P H Y S I C S

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN AND CO., LIMITED
LONDON · BOMBAY · CALCUTTA · MADRAS
MELBOURNE

THE MACMILLAN COMPANY
OF CANADA, LIMITED
TORONTO

PHYSICS

BY FRANK L. ROBESON

Professor of Physics, Virginia Polytechnic Institute

THE MACMILLAN COMPANY

NEW YORK · 1943

Printed in the United States of America
Published April, 1942; reprinted October, 1942; January, June, July, August, 1943.

PREFACE

For years the author has cherished the thought that it should be possible to have a textbook of physics in which the subject matter was presented in such a clear and orderly fashion that it would not be necessary for the teacher to explain the text, and that in consequence his time in the class room could be devoted to demonstrations, discussions, and the solution of problems. The present volume is an attempt to achieve that ideal.

In many instances, classes in general physics are composed of students who have had no previous training in the subject, as well as of those who have had an introductory course. The treatment here given is designed to meet the requirements of both groups.

No apology is made for the use of elementary mathematics which is now either required for entrance or taken concurrently with physics in all standard colleges. Nothing more alarming than the law of cosines is needed. In a few instances calculus is employed, but that may be omitted without loss of continuity. Judicious repetition and liberal cross-references are employed to fix the principles more firmly in the readers' minds; and, for the serious students, many footnotes indicate where fuller discussions may be found. Historical sidelights are interspersed to maintain the chronological perspective and to make the prescription more palatable.

The author wishes to acknowledge his indebtedness to his co-workers in the Department of Physics at the Virginia Polytechnic Institute whose unfailing interest and assistance have made this book possible, to the late Professor M. Houston Eoff who read all the proof, to Professor G. A. Van Lear of the University of Oklahoma and to Professor A. B. Cardwell of Kansas State College for many helpful suggestions, and to the various publishers and industrial firms who without exception gave permission for the use of illustrative material.

Although great care has been exercised to have the book free from textual errors, it is too much to hope that all have been eliminated; and the writer will appreciate having his attention called to those that remain.

F. L. R.

BLACKSBURG, VIRGINIA
March 1, 1942

CONTENTS

INTRODUCTION

MECHANICS

HEAT

ELECTRICITY AND MAGNETISM

ELECTRODYNAMICS

LIGHT

SYMBOLS

$\equiv$	means "is identical," or "is defined as"
$\neq$	means "is not equal"
$\doteq$	means "is approximately equal," or "approaches"
$\propto$	means "varies as," or "is proportional to"
$\bar{x}$	means "average value," or "mean value"
Σ	means algebraic sum of

GREEK ALPHABET

A	α	Alpha	N	ν	Nu
B	β	Beta	Ξ	ξ	Xi
Γ	γ	Gamma	O	o	Omicron
Δ	δ	Delta	Π	π	Pi
E	ϵ	Epsilon	P	ρ	Rho
Z	ζ	Zeta	Σ	σ	Sigma
H	η	Eta	T	τ	Tau
Θ	θ	Theta	Υ	υ	Upsilon
I	ι	Iota	Φ	ϕ	Phi
K	κ	Kappa	X	χ	Chi
Λ	λ	Lambda	Ψ	ψ	Psi
M	μ	Mu	Ω	ω	Omega

INTRODUCTION

Read, not to contradict and confute, nor to believe and take for granted, nor to find talk and discourse, but to weigh and consider.

—Lord Bacon

[illegible]

CHAPTER I

DEFINITIONS

1. The science of physics is so intimately involved in everyday affairs that without it our present civilization would be inconceivable. Clocks, telephones, airplanes, and talking pictures are familiar examples of its innumerable applications. No other single science contributes more to the comfort and compass of modern life.

By the term **science** (from Latin, *scire* = to know) we mean systematized knowledge. Such knowledge of the objects and processes occurring in nature—**natural science**—has two main divisions, namely, biological and physical.

The biological sciences treat of living things; the physical sciences, of inanimate things. Fundamental among the physical sciences are physics and chemistry. In fact, these two and the abstract science of mathematics are indispensable in all the natural sciences. Astronomy, geology, navigation, and engineering, for example, are essentially physics, mathematics, and chemistry; while medicine and agriculture involve biology also.

2. Matter and energy. The objects that we perceive—trees, rocks, stars, air, etc.—are called bodies. The material, or stuff, of which these bodies are made is called matter.

We frequently observe matter changing its position, its condition, or even its kind: a train moving, water boiling, wood burning. Any change that matter undergoes we consider to be due to something which we call energy.

More precisely defined:

Matter is anything that offers resistance to change in its condition of rest or motion.

Energy is that which causes a change in matter.

These two, matter and energy, are considered by many scientists to be the sole entities in the physical world. There is evidence, however, that matter and energy are mutually convertible.

3. Physics and chemistry are the science of matter and energy.

It is becoming increasingly difficult to distinguish between physics and chemistry. For the present, physics may be considered to deal with those phenomena in which the kind of matter does not change; and chemistry, with those in which the kind of matter does change.

Thus when coal is hauled from a mine it remains the same coal even though it has been moved and perhaps broken. The transaction is one of physics. But when coal is burned, some of it passes off as carbon dioxide and water and some remains as ash: the coal has been changed into other kinds of matter. This latter change belongs to chemistry.

Later on, however, it will be seen that often this distinction between physics and chemistry is not maintained.

4. The scientific method. The method of science consists in observation, investigation, and explanation of the phenomena, or occurrences, in nature. When the materials and circumstances essential to the occurrence have been found and set in order so that the phenomenon may be reproduced at will,* and the whole transaction has been described accurately, we then say that we have the law of that phenomenon.

A physical law, or principle, is a statement by which we can predict the effect of a given cause.

The first postulate of science affirms that the same cause always produces the same effect. Science is based so completely on this belief that when causes which seem to be the same produce different results, the causes are re-examined; and invariably it has been found that they were not the same. There have been many instances in which this has led to important discoveries. For example, the great English physicist, Lord Rayleigh, in preparing pure nitrogen, found that nitrogen obtained from air was distinctly heavier than that obtained by decomposing ammonium nitrate. He assumed, therefore, that nitrogen was not the "cause" of the weight in both cases. Further investigation showed that the nitrogen obtained from the air was mixed with a gas, hitherto unknown. Thus the important gas argon (Greek, lazy) was discovered.

The constant endeavor of scientists is to generalize; that is, to group together as many laws as possible into one general law, thereby reducing the number of laws to a minimum. Thus the

* There are, of course, many cases when this cannot be carried out literally.

laws governing the fall of an apple, the rise of a balloon, and the motion of the planets may be shown to be special cases of one general law of gravitation.

Experiment is essential to the progress of science. An **experiment** is a procedure planned for a more or less definite purpose. The facts of science are often discovered or suspected by a casual experience, or by deduction from previously established principles. But in every case final acceptance of a theory is withheld until by experiment its truth has been demonstrated beyond reasonable doubt.

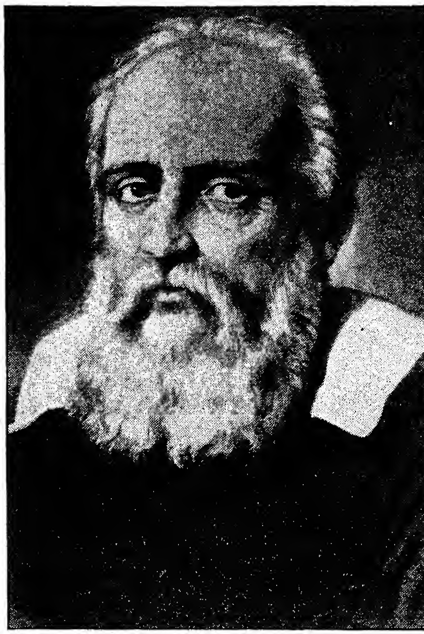


FIG. 1. Galileo Galilei

Up to the time of Galileo (1564–1642), statements were accepted as truth on authority. Aristotle (384–322 B.C.), who as late as the sixteenth century was still the leading authority in scientific matters, had stated that heavy bodies fall faster than light ones. So great was his prestige that no one had questioned the statement in 1900 years, nor had any attempt been made to test it. But Galileo was an experimenter; according to the legend, he tried dropping stones of various weights from the top of the Leaning Tower of Pisa (Fig. 2).* Contrary to the statement of Aristotle, they all fell to the ground in the same time. Thus the long established spell of Aristotle was broken, and Galileo is now recognized as the father of experimental science.

A phenomenon is said to be **explained** when it is shown to be in accord with well established laws.

Statements such as, "The moon remains in its orbit by the law

* In *Aristotle, Galileo, and the Tower of Pisa* (Ithaca, Cornell University Press, 1934), Prof. Lane Cooper seriously questions whether this experiment was ever actually made.

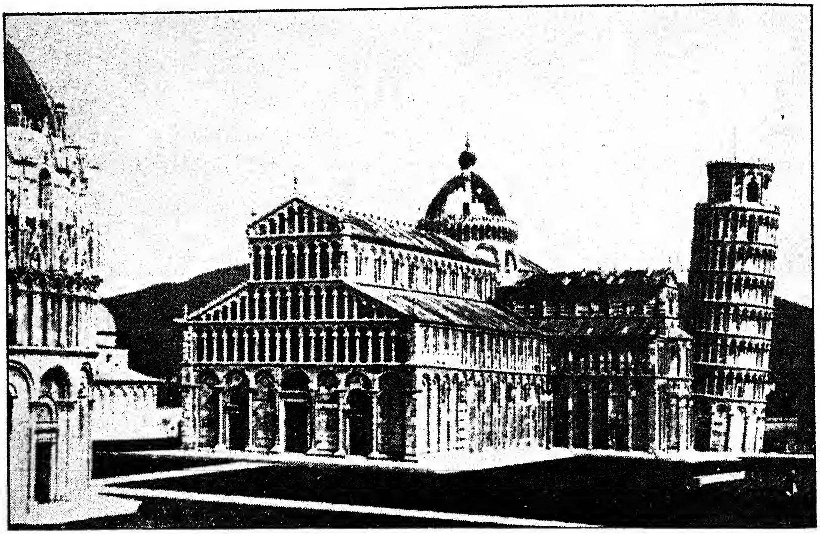


FIG. 2. Cathedral and Leaning Tower of Pisa. (Courtesy of Mr. Thomas S. Derr)

of gravitation," are often heard. It should be borne in mind that occurrences do not take place in nature on account of the laws, but **in accordance with the laws.** The laws, or principles, are in fact merely statements of the manner in which phenomena are observed to occur.

5. States of matter. From everyday observation we recognize matter in three states: solid, liquid, and gas.

Solids have definite shape and size of their own and resist any attempt to change their shape or size; for example, ice.

Liquids have definite size (volume) but take the shape of the containing vessel. They show a **free upper surface;** for example, water.

Gases have neither size nor shape of their own, but take both from the containing vessel. They show no free upper surface; for example, steam.

The term **fluid** includes both liquids and gases, because both flow.

CHAPTER II

MEASUREMENTS

6. The term **magnitude, or quantity,** is used in an abstract way to mean something that is susceptible of measurement, such as length, volume, speed, and temperature. Magnitude and quantity are used also to designate how much of a thing there is; for example, the magnitude, or quantity, of a pile of earth may be 50 cubic feet.

Measurement consists in determining how many times a given magnitude contains another magnitude of the same kind which has been chosen as a *unit.* Thus we choose the foot as a unit of length, and measure the length of a table by finding how many times it contains the unit—perhaps 5 times. The length is then said to be 5 feet. It is not always possible to divide the quantity into parts each the size of the unit. We cannot, for example, divide a temperature into pieces of temperature each equal to the unit. Measurements of such quantities have to be made by indirect methods.

The expression for the length of the table is seen to consist of two parts: the **numeric** (5) and the **unit** (foot). Such numbers are called **compound numbers.** A number such as 5, occurring without a unit, is called a *pure*, or *abstract*, *number*. The trigonometric functions and angular measures are all pure numbers (see Sec. 10).

7. Fundamental magnitudes, or quantities. Numerous as are the quantities to be measured that occur in nature, we shall see as we proceed that, with a few exceptions, they can all be measured in terms of length, mass, and time.

Length (L), mass (M), and time (T) are therefore called fundamental quantities, because practically all other quantities may be expressed in terms of them. Many other selections of fundamental units may be made. Length, mass, and time are merely the simplest choice at first.

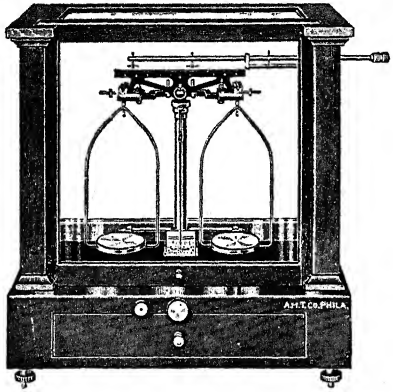


FIG. 3. Beam Balance

Length and time are scientifically indefinable inasmuch as they cannot be described in more familiar terms.

Mass may be defined for the present as **the amount of matter in a body as determined by weighing** * **on a beam balance** (Fig. 3). (For distinction between mass and weight, see Sec. 44.)

8. Fundamental units. For each of the fundamental quantities a unit is chosen; and since there are two systems of units in common use—the metric and the British—we have two sets of fundamental units:

FUNDAMENTAL UNITS

Quantity	Units	
	Metric (cgs)	British (fps)
Length	Centimeter	Foot
Mass	Gram	Pound
Time	Second	Second

The Giorgi (M K S) system. In 1901, Professor G. Giorgi of Rome pointed out that by taking the meter, the kilogram, and the second as fundamental units, the absolute electromagnetic units (Sec. 380) take the same values as the present practical electrical units (Sec. 395). The International Committee on Weights and Measures has recommended the adoption of this system.

These fundamental units are defined in terms of certain standards adopted by legislative enactment and are internationally recognized.

The **standard of length** is the International Prototype Meter kept at the International Bureau of Weights and Measures near

* Unfortunately, in English the verb *to mass* is used only in the sense of "to assemble" (as of troops). Hence we use the verb *to weigh* in two senses: (1) to determine mass, and (2) to determine weight.

Paris. This is a bar of platinum (90%) and iridium (10%), having two fine transverse lines drawn on its neutral surface near the ends of the bar (Fig. 4). The meter is defined as the distance between the centers of these two lines when the bar is at the temperature of melting ice.

The meter was originally intended to be 1/40,000,000 of the meridian through Paris. This value was computed from measurements of an arc of the meridian between Barcelona and Dunkirk made by a commission appointed by the French government (1791). More precise measurement has shown that the length of the meridian is more nearly 40,008,400 meters.

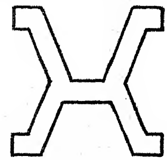


FIG. 4. Cross Section of Standard Meter (Actual Size)

In 1894, at the request of the French government, Michelson determined the length of the standard meter at 15°C as 1,553,163.5 wave lengths of the red line in the spectrum of cadmium. The

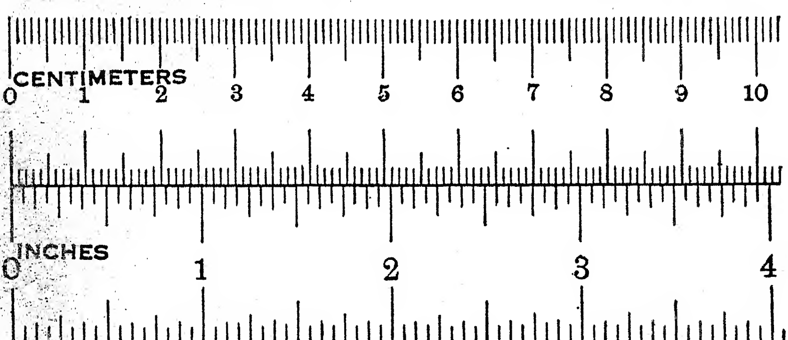


FIG. 5. Scale of Centimeters and Inches

ultimate standard is, therefore, the wave length of the red line of the spectrum of cadmium, which is assumed to be invariable.

$$1 \text{ centimeter (cm)} \equiv \tfrac{1}{100} \text{ meter.}$$

The **standard yard** was originally the distance between two marks on a bar of bronze in the possession of the British government. The yard is now defined in the United States as 3600/3937 of the standard meter.

$$1 \text{ foot} \equiv \tfrac{1}{3} \text{ yard}$$
$$1 \text{ inch} \equiv \tfrac{1}{12} \text{ foot.}$$

The meaning of the prefixes of the metric system should be memorized.

Mega-meter =	1,000,000 meters
Kilo-meter =	1000 meters
Hecto-meter =	100 meters
Deka-meter =	10 meters
Deci-meter =	1/10 meter
Centi-meter =	1/100 meter
Milli-meter =	1/1000 meter
Micro-meter =	1/1,000,000 meter.

These prefixes may also be applied to the other metric units; for instance, milliliter, kilogram, etc.

The **standard of mass** is the International Prototype Kilogram. This is a cylinder of platinum kept by the French government at Sèvres, near Paris.

It was intended that this mass should be equal to the mass of one cubic decimeter (1000 cubic centimeters) of water at 3.98°C, at which temperature water has its maximum density. On account of slight errors in making the prototype kilogram, this relation is not quite true; hence the kilogram of mass, like the meter, is an arbitrary unit.

1 gram (gm) $\equiv \frac{1}{1000}$ kilogram (kg)
= mass of 1 milliliter (ml) of water at 3.98°C
= mass of 1 cubic centimeter (cm^3) of water (approximately)

The **standard pound** was at first taken arbitrarily as the mass of a cylinder of platinum kept by the British government. The pound of mass is now defined in the United States as 1/2.20462 kilogram.

The **standard of time** is the mean, or average, solar day; i.e., the average time required for the earth to make one revolution on its axis with reference to the center of the sun. Hence, **the earth is our primary clock.**

Astronomers use the sidereal day, which is the time required for the earth to make one revolution on its axis with reference to the stars.

$$1 \text{ year} \equiv 365\tfrac{1}{4} \text{ mean solar days} = 366\tfrac{1}{4} \text{ sidereal days}$$

$$\textbf{1 mean solar second} \equiv \frac{1}{86{,}400} \text{ mean solar day.}$$

9. Derived units. Practically all other units are derived from the fundamental units.

(*a*) *Units of area.* Consider the rectangle of Fig. 6:

$$\text{Area} = \text{length} \times \text{width}.$$

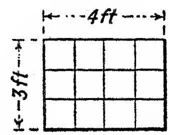


Fig. 6. Dimensions of Area

But width is another length measured crosswise; therefore,

$$\begin{aligned}\text{Area} &= \text{length} \times \text{length}\\ &= 4 \text{ ft} \times 3 \text{ ft} = 12 \text{ ft}^2.\end{aligned}$$

Here the unit of area is the **square foot;** i.e., the area of a square whose side is one foot long.

Instead of using the foot as the unit of length in the above case, let [L] represent any unit of length. The more general expression for the rectangle is then

$$\text{Area} = 4\,[L] \times 3\,[L] = 12\,[L^2],$$

which shows that the unit of area always is logically the area of a square whose side is the unit of length that is being used.

Hence units of area are the square centimeter (cm^2), the square foot (ft^2), etc.

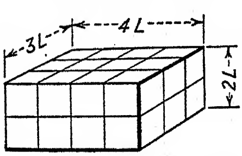


Fig. 7. Dimensions of Volume

(*b*) *Units of volume.* Similarly for a rectangular parallelopiped (Fig. 7), where [L] is any unit of length,

$$\text{Volume} = 4\,[L] \times 3\,[L] \times 2\,[L] = 24\,[L^3].$$

Thus we see that, for any case, the unit of volume may be derived as the volume of a cube whose edge is the unit of length.

The commonly used units of volume are the cubic centimeter (cm^3), the cubic foot (ft^3), etc.

The liter. The prototype kilogram was intended to be exactly the mass of 1 cubic decimeter of water at 3.98°C. As made, it is slightly too large. Hence the unit of volume in the metric system is defined as follows:

1 liter $\equiv$ volume of water having mass of 1 kg at its temperature of maximum density
$= 1.000027$ cubic decimeter (dm^3)

so that

$$1 \text{ milliliter} \equiv \tfrac{1}{1000} \text{ liter} = 1.000027 \text{ cm}^3.$$

(*c*) *Units of speed.* If a person moves 40 feet in 5 seconds, we say his average speed is 8 feet per second:

$$\text{Speed} = \frac{40 \text{ ft}}{5 \text{ sec}} = 8 \frac{\text{ft}}{\text{sec}}.$$

Thus the unit of speed, 1 foot per second (ft/sec), is derived from the fundamental units of length and time.

Or if, as before, we let $[L]$ and $[T]$ represent any units of length and time, respectively, we have:

$$\text{Speed} = \frac{40}{5}\frac{[L]}{[T]} = 8\left[\frac{L}{T}\right] = 8\,[LT^{-1}],$$

where $\left[\frac{L}{T}\right]$, or $[LT^{-1}]$, may represent any unit of speed; e.g., 1 cm/sec, 1 mi/hr, etc., in different problems.

(*d*). *Units of density.* **Density D is defined as mass M per unit of volume V.** Algebraically,

$$D \equiv \frac{M}{V}.$$

In the case of water, the mass is 1 gram when the volume is 1 milliliter, for that is the definition of the milliliter.

Therefore, **for water,**

$$D = \frac{1 \text{ gm}}{1 \text{ ml}} = 1 \frac{\text{gm}}{\text{ml}} \doteq 1 \frac{\text{gm}}{\text{cm}^3}.$$

Similarly, in the fps system, 1 cubic foot of water has a mass of 62.4 pounds, so that again, **for water,**

$$D = \frac{62.4 \text{ lb}}{1 \text{ ft}^3} = 62.4 \frac{\text{lb}}{\text{ft}^3}.$$

In general, if $[M]$ is any unit of mass and $[V]$ is any unit of volume, the mass of any body would be $a[M]$ and its volume

would be $b[V]$, where a and b are necessary numerical coefficients. For any body, then,

$$D = \frac{a\,[M]}{b\,[V]} = \frac{a}{b}\left[\frac{M}{L^3}\right] = \frac{a}{b}\,[ML^{-3}],$$

where a/b is the numeric and $[ML^{-3}]$ is the unit for any density whatever.

10. Dimensional formulas. An algebraic expression which shows how the derived unit of a quantity is made up from the fundamental units is called the *dimensional formula* of that quantity.

Thus, in the preceding section, we obtained the dimensional formulas for the following quantities:

Area	$[L^2]$
Volume	$[L^3]$
Speed	$[LT^{-1}]$
Density	$[ML^{-3}]$.

The brackets mean that the enclosed letters refer to the kind of unit and not to its magnitude.

An abstract, or pure, number such as 5 means five of the pure number 1 and does not refer to any particular unit such as the foot, the kilogram, or the liter. From algebra we know that any quantity to the zero power equals 1; thus $x^0 = 1$; $L^0 = 1$; $M^0 = 1$; $T^0 = 1$. Hence the dimensional formula of a pure number is $[L^0M^0T^0] = [1]$.

Since the trigonometric functions of an angle are defined as the ratios of certain sides in a right triangle which contains that angle, we have (Fig. 8):

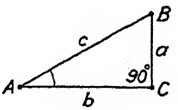


FIG. 8

$$\sin A \equiv \frac{\text{opposite side}}{\text{hypotenuse}} = \frac{a}{c} = \frac{[L]}{[L]} = [L^0] = [1].$$

and similarly for the other trigonometric functions.

Also, as a physical quantity, angle is defined (Sec. 11) as

$$\text{angle} \equiv \frac{\text{arc}}{\text{radius}} = \frac{[L]}{[L]} = [L^0] = [1].$$

Hence, **all the trigonometric functions,** and **angles** in circular measure, **have the dimensional formula of pure numbers.**

Dimensional formulas are often very useful both in numerical computations and in theoretical reasoning, which will become evident as we proceed.

11. Circular measure of angles. It will be recalled from trigonometry that there are two principal ways of measuring angles:

1. The sexagesimal system, in which the units are degrees, minutes, and seconds.

2. The circular system, in which the unit angle is the **radian.** In the latter system, angle is defined as follows:

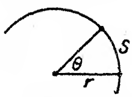


FIG. 9

$$\text{Angle} \equiv \frac{\text{Intercepted Arc}}{\text{Radius}} \tag{1}$$

or in symbols, as shown in Fig. 9,

$$\theta \text{ radians} = \frac{s\,\cancel{\text{cm}}}{r\,\cancel{\text{cm}}}. \tag{2}$$

From Eq. (2) it is seen that $\theta = 1$ when arc s = radius r; i.e., **a radian is the angle subtended at the center of a circle by an arc equal to the radius.** The abbreviation for the word "radian" is **rdn.**

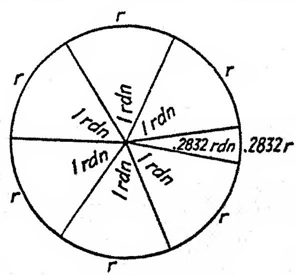


FIG. 10

From geometry we know that a circumference = $2\pi r$. Hence we can lay off the radius along the circumference 2π times, as in Fig. 10, and each of these lengths equal to the radius will subtend an angle of 1 radian at the center. Therefore,

$$\begin{aligned} 2\pi \text{ rdn} &= 360° \\ 1 \text{ rdn} &= \frac{360°}{2\pi} = \frac{180°}{\pi} \\ &= 57°17'45''. \end{aligned} \tag{3}$$

This last value is seldom needed, however.

An advantage of the circular system over the sexagesimal system is that it is possible at once, on knowing an angle, to obtain the length of its intercepted arc on any circle by multiplying the value of the angle in radians by the length of the radius of the circle, as is shown by clearing fractions in Eq. (2):

$$s = r\theta. \tag{4}$$

Thus, if an angle is 2.3 rdn (Fig. 11), the intercepted arc length on the 2-cm circle will be 4.6 cm; and on the 5-cm circle, 11.5 cm.

12. Conversion factors. The following conversion factors are used so frequently that they should be memorized.

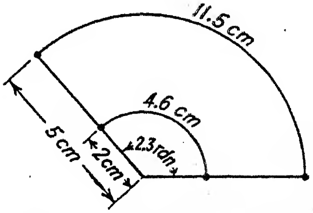


FIG. 11. Advantage of Circular Measure of Angle

1 inch = 2.54 centimeters (more accurately = 2.54001)
1 kilogram = 2.20 pounds (more accurately = 2.20462)
1 mile = 5280 feet
1 gallon = 231 cubic inches.

Solved Problems

1. Compute the number of feet in 1 kilometer.

Known:

$$1 \text{ km} = 1000 \text{ m} = 100{,}000 \text{ cm}$$
$$1 \text{ in.} = 2.5400 \text{ cm}$$
$$1 \text{ ft} = 12 \text{ in.}$$

Solution:

$$1 \text{ km} = 100{,}000 \text{ cm} = \frac{100{,}000 \cancel{\text{cm}}}{2.54 \frac{\cancel{\text{cm}}}{\text{in.}}} = 39{,}370 \text{ in.}$$

$$= \frac{39{,}370 \cancel{\text{in.}}}{12 \frac{\cancel{\text{in.}}}{\text{ft}}} = 3281 \text{ ft.}$$

2. Compute the number of cubic centimeters in 1 cubic foot.

Known:

$$1 \text{ ft} = 12 \text{ in.}$$
$$1 \text{ in.} = 2.5400 \text{ cm.}$$

Solution:

$$1 \text{ ft} = 12 \cancel{\text{in.}} \times 2.5400 \frac{\text{cm}}{\cancel{\text{in.}}} = 30.480 \text{ cm}$$
$$1 \text{ ft}^3 = (30.480 \text{ cm})^3 = 28{,}317 \text{ cm}^3.$$

3. Compute the number of kilograms in 1 ton.

Known:

$$1 \text{ ton} = 2000 \text{ lb}$$
$$1 \text{ kg} = 2.2046 \text{ lb.}$$

Solution:

$$1 \text{ ton} = \frac{2000 \cancel{\text{lb}}}{2.2046 \frac{\cancel{\text{lb}}}{\text{kg}}} = 907.19 \text{ kg.}$$

4. Compute the mass of 1 cubic inch of water at 4°C.

Known:

1 in. = 2.5400 cm
1 ml = 1.000027 cm^3
1 ml of water at 3.98°C has a mass of 1 gm by definition.

Solution:

$$1 \text{ in.}^3 = (2.54 \text{ cm})^3 = 16.387 \text{ cm}^3$$
$$= \frac{16.387 \cancel{\text{cm}^3}}{1.000027 \frac{\cancel{\text{cm}^3}}{\text{ml}}} = 16.386 \text{ ml.}$$

But

1 ml of water at 3.98°C has a mass of 1 gm.

Therefore

1 $in.^3$ of water at 3.98°C has a mass of 16.386 gm.

PROBLEMS

1. From the relation 1 in. = 2.54 cm, compute the number of feet in 1 meter.

2. From the relation 1 in. = 2.54 cm, compute the number of miles in 1 kilometer.

3. From the relation 1 in. = 2.54 cm, compute the number of yards in the 100-meter dash.

4. Compute the number of cubic centimeters in 1 in^3.

5. If 1 kg = 2.2046 lb, compute the number of grams in 1 lb.

6. Compute the mass of 1 gal of water, from the mass of 1 ft^3.

7. Calculate the number of quarts in 1 liter.

8. From the definition of a gram, compute the mass of 1 ft^3 of water.

9. If a milliliter of water weighs 1 gm, what is the weight of 1 gal of water?

10. Calculate the number of milliliters in a fluid ounce. (16 fluid oz = 1 pint)

11. Compute the mass of a column of water 1 ft high and 1 $in.^2$ in cross section.

12. Express 60° in radians. Express 0.8 rdn in degrees.

13. An arc of 3.6 in. subtends what central angle in a circle of 2-in. radius? Of 5-in. radius?

14. What arc is intercepted on a circle of 8-cm radius by a central angle of 3 rdn? Of 3°?

MECHANICS

When you can measure what you are speaking about and express it in numbers, you know something about it; and when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind. It may be the beginning of knowledge, but you have scarcely in your thought advanced to the stage of a science.

—Lord Kelvin

CHAPTER III

KINEMATICS

13. Mechanics is that branch of physics that treats of motion, rest being regarded as a special case of motion. It is divided into **kinematics,** which is the science of pure motion without regard to its cause, and **dynamics,** which deals with the causes and results of motion.

14. Motion. When a body is changing its position, it is said to have *motion.* To describe a body's motion, it is necessary to locate the positions of its several points. This is done by various

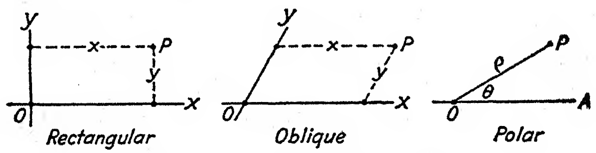


FIG. 12. Systems of Coordinates

systems of coordinates, the following—as in analytic geometry—being the most common (Fig. 12).

(a) In **rectangular** coordinates a point's position in a plane is indicated by giving its distances, x and y, from two fixed perpendicular lines called the y-axis and x-axis, respectively.

(b) If the axes are not at right angles the coordinates are called *oblique*, the distances being measured parallel to the axes as shown.

(c) In the **polar** system of coordinates, a position is designated by giving the length of a **radius vector** (ρ) and the angle (θ) which the radius vector makes with a certain line OA called the **initial line.**

In the first two systems the point O is called the **origin**; in the last, the **pole.**

For representing the position of a point in space, three planes are taken, making known angles with one another, and a corre-

sponding third coordinate is given.* Any such set of points, lines, or planes used as a base from which to locate other points is called a **frame of reference.** The frame of reference may be chosen so as to make each problem as simple as possible.

Since the position of a point, and hence its changes of position, are necessarily determined by measurements from, **or relative to,** the parts of the frame of reference, it is seen that **all motion is relative.** We ordinarily consider the frame of reference to be fixed, i.e., to have no motion relative to the observer. Both the frame of reference and the observer, however, may have motion with reference to some other frame of reference. For example, a person standing at the intersection of two city streets observes a car moving down one of those streets at 25 mph (mi/hr). This is its speed **relative** only to the observer, who is considered to be at rest. But it is well known that the earth is moving with reference to the sun; hence the motion of the car with reference to the sun would be a combination of those two motions. The solar system in turn moves with reference to the stars, and so this idea of the relativity of motion may be extended indefinitely.

15. Particle. A particle is defined as a piece of matter so small that its linear dimensions may be disregarded in the question under discussion. We then have to consider only its position and its mass. For many cases in physics we may think of a particle as being a molecule as it is defined in chemistry. In mathematical discussions the particle has the infinitesimal dimensions dx, dy, dz, or their equivalents.

A **rigid body** is a body in which the distance between each pair of points remains constant. No perfectly rigid body exists, but for simplicity all bodies are considered to be rigid unless otherwise stated.

16. Types of motion. Two kinds of motion are recognized:

(1.) **Translation,** in which all particles of the body move in parallel paths with the same linear speed.

(2.) **Rotation,** in which all particles describe circles in parallel planes about the same axis and with the same angular speed.

Any motion, however complicated, may be shown to be equiva-

* See any text on analytic geometry.

lent to a combination of translation and rotation, provided the body is rigid.

The movement of the cage of an elevator in its guides is an example of translation; that of an electric fan, of rotation. The motion of the wheels of a railway car is easily seen to consist of both translation and rotation.

Motion in a straight line is called **rectilinear** motion. Motion not rectilinear is called **curvilinear.**

17. Displacement. If a particle has moved from A to B (Fig. 13), its displacement is defined in magnitude and direction by the straight line AB, regardless of whether the particle actually followed the straight line or pursued some other path, such as the dotted one. The length of any line other than the straight line between two points should be called a **distance** between the points **and not** the displacement.

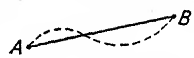


FIG. 13. Linear Displacement

Rotational, or angular, displacement is the angle through which a body has turned with a given direction of the axis. Thus (Fig. 14), if AB is the position of any line on the body before the motion, and $A'B'$ the position of the same line after the motion, the angle θ between AB and $A'B'$ is the angular displacement.

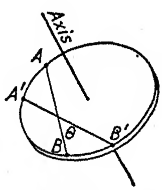


FIG. 14. Angular Displacement

18. Velocity is the time rate of change of displacement. Since displacement has direction, velocity also has direction.

The term **speed** is used for the time rate of motion without regard to direction; i.e., speed is the magnitude only of velocity.

When a body moves so that its displacements in any two equal intervals of time are equal, it is said to have **uniform motion.** Its velocity is then constant,* so that, algebraically:

Uniform motion is defined by $v =$ constant, for translation
or $\omega =$ constant, for rotation.

Motion is seldom uniform, however; hence we define **average speed** as the total distance divided by the total time.

* In uniform motion, the direction as well as the magnitude must be constant.

Let x_1 and x_2 (Fig. 15) be the distances of a particle on the line OX from point O at clock times t_1 and t_2, respectively. Then

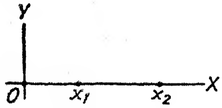


Fig. 15

$$\text{magnitude of distance} = x_2 - x_1$$
$$\text{and total time} = t_2 - t_1$$

and the above definition may be written in the shorter and more convenient form:

$$\text{Average speed} \quad \bar{v} \equiv \frac{x_2 - x_1}{t_2 - t_1}. \tag{5}$$

Thus, suppose an automobile traveling northward along a north-and-south line is 5 miles from the garage G Fig. 16 at 9:10 A.M. and 25 miles from it at 9:40 A.M. Then,

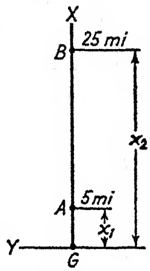


Fig. 16

$$\begin{aligned} \text{Total distance} &= x_2 - x_1 = 25 - 5 = 20 \text{ mi} \\ \text{Total time} &= t_2 - t_1 = 9{:}40 - 9{:}10 = 30 \text{ min} \\ &= 0.5 \text{ hr} \end{aligned}$$

so that

$$\text{Average speed} \quad \bar{v} \equiv \frac{x_2 - x_1}{t_2 - t_1} = \frac{20 \text{ mi}}{0.5 \text{ hr}} = 40 \text{ mi/hr}$$

and

$$\text{Average velocity} \quad \bar{v} = 40 \text{ mi/hr, } \textbf{northward.}$$

There is nothing in Eq. (5) to designate direction: that will be taken up in Sec. 20. Hardly any road is straight, but what has been said applies to curvilinear motion as well as to rectilinear; hence the word **distance** was used and not **displacement.**

It should be noted that the value 40 mi/hr is the average speed for the whole trip, and not the speed at any particular instant; for the car actually starts from rest ($v = 0$), speeds up to a maximum value, and may slow down to rest again at B.

Instantaneous velocity. If we wish to determine the speed and velocity at any particular instant, instead of the average value, we locate the point P (Fig. 17) on its path where the car was at that instant, and determine (say, with a stop watch) the time $(t_2' - t_1')$ required for the car to traverse a very short distance $(x_2' - x_1')$ having P at its cen-

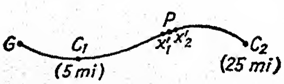


Fig. 17

ter—so short that during this interval the speed may be considered constant.

Such a very short distance and interval of time are represented by the symbols Δx and Δt, respectively. Hence, for practical purposes,

$$\text{Instantaneous speed} \equiv \frac{x_2' - x_1'}{t_2' - t_1'} \equiv \frac{\Delta x}{\Delta t}. \tag{6}$$

The instantaneous velocity v has the numerical value of the speed and the direction of the little segment $x_1'x_2'$; i.e., it has the direction of the tangent at P.

In calculus it is shown rigorously that the ratio $\Delta x/\Delta t$ approaches in many cases a definite limiting value as Δt approaches 0. This limiting value of $\Delta x/\Delta t$ is called the first derivative of x with respect to t, and is written dx/dt, so that

$$\text{Instantaneous} \quad v \equiv \left[\frac{\Delta x}{\Delta t}\right]_{\lim \Delta t \doteq 0} \equiv \frac{dx}{dt}. \tag{7}$$

Rotation is treated in an exactly similar way. Let θ_1 and θ_2 (Fig. 18) be the angles through which a body has turned at clock times t_1 and t_2, respectively. Then,

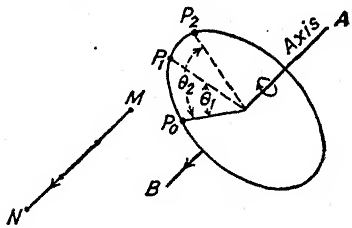


FIG. 18

Magnitude of angular motion $= \theta_2 - \theta_1$

Time of angular motion $= t_2 - t_1$.

Therefore,

$$\text{Average angular speed,} \quad \bar{\omega} \equiv \frac{\theta_2 - \theta_1}{t_2 - t_1}. \tag{8}$$

If the rotation is uniform, this average speed is also the speed at any instant. Otherwise, we must divide the small angle $\Delta\theta$ by the short time Δt in which the body turns through $\Delta\theta$, the Δt being taken so short that during that time the angular speed may be considered constant. Then,

$$\text{Instantaneous angular speed,} \ \omega \equiv \left[\frac{\Delta \theta}{\Delta t}\right]_{\lim \Delta t \doteq 0} \equiv \frac{d\theta}{dt}. \tag{9}$$

The instantaneous angular velocity ω has the magnitude given by Eq. (9). To this must be added a statement of the direction of

rotation, which is indicated by the direction of the axis as described in Sec. 23.

For example, consider the flywheel of an engine which at 9:35 A.M. has turned through 200 revolutions from some initial position, and which at 9:39 A.M. has turned through 1000 revolutions from the same initial position, the axis of rotation retaining its initial direction. Then,

Total angular displacement = 800 revolutions
Total time required = 4 minutes

so that

$$\text{Average angular speed} = \frac{1000 - 200}{9{:}39 - 9{:}35} = \frac{800 \text{ rev}}{4 \text{ min}} = 200 \text{ rev/min.}$$

If the direction of the axis had changed, that must be noted separately for the present.

19. Acceleration. Acceleration (a) is the time rate of change of velocity (v).

Since velocity has direction, acceleration also has direction. In algebraic shorthand, average acceleration is defined by the following equations, provided the direction is constant:

$$\bar{a} \equiv \frac{v_2 - v_1}{t_2 - t_1} \quad \text{for translation} \tag{10}$$

and

$$\bar{\alpha} \equiv \frac{\omega_2 - \omega_1}{t_2 - t_1} \quad \text{for rotation} \tag{11}$$

where v_2, ω_2 and v_1, ω_1 are the linear and angular velocities at the clock times t_2 and t_1, respectively.

In a more advanced course, Vector Analysis, it will be found that Eqs. (10) and (11), as well as (5), (7), (8), and (9), may be construed to include the element of direction; but in this work we shall employ them with reference to magnitude only, and consider the question of direction independently.

Acceleration may be either positive or negative. Negative acceleration is often called **deceleration.** Thus, an automobile has two accelerators: the one for positive acceleration is called the "accelerator"; the one for negative, the "brake." **When a body is accelerated, it is always either speeding up or slowing**

down, except in the case of a body having uniform circular motion (see Sec. 33).

If the car in Fig. 16 had a speed of 35 mph at A and 45 mph at B, then

Total change of speed $= 45 - 35 = 10$ mi/hr
Total time required $= 9{:}40 - 9{:}10 = 30$ min $= 0.5$ hr.

Therefore,

$$\text{Average acceleration, } \bar{a} \equiv \frac{v_2 - v_1}{t_2 - t_1} = \frac{10 \text{ mi/hr}}{0.5 \text{ hr}}$$
$$= 20 \text{ mi/hr in 1 hr}$$
$$= 20 \text{ mi/hr per hour}$$
$$= 20 \text{ mi/hr}^2$$

and the direction, **northward**, must be stated.

Notice that since velocity has the unit of time in the denominator once, and since acceleration is found by dividing change of velocity by time, acceleration will always have the unit of time as a factor in the denominator twice; i.e., **the dimensional formula for acceleration is** $[LT^{-2}]$.

Instantaneous acceleration. If the acceleration that the car has just as it passes a point P (Fig. 17) is desired, we must take its speed (from its speedometer, say) at two points having P between them and so close together that the rate of change of velocity between these two points may be considered constant. For an extremely short element, or piece, of the path like this the change of velocity $(v_2 - v_1)$ is called Δv, and the corresponding time $(t_2 - t_1)$, Δt. Hence the instantaneous acceleration of the car as it passes point P is:

$$a \equiv \frac{\Delta v}{\Delta t}, \text{ or in the limit as } \quad \Delta t \doteq 0, \quad a \equiv \frac{dv}{dt}. \tag{12}$$

Similarly, instantaneous angular acceleration is:

$$\alpha \equiv \frac{\Delta \omega}{\Delta t}, \text{ or in the limit as } \quad \Delta t \doteq 0, \quad \alpha \equiv \frac{d\omega}{dt}. \tag{13}$$

Acceleration due to gravity. The most common acceleration which we encounter is that of freely falling bodies. We call this the acceleration of gravity, and represent it by the symbol g. At

Washington, D.C., it has the value 980.1 cm/sec^2 (= 32.16 ft/sec^2), but it differs slightly from this value at different places.

These values are for free fall in a **vacuum**, thus taking no account of air friction, which increases when the speed increases. Actually, after a dummy dropped from an airplane has fallen about 1600 ft, the resistance of the air is equal to the force of gravity so that the acceleration is zero. The dummy then continues to fall with a uniform speed of about 120 mi/hr.

20. Scalars and vectors. It will have been observed by now that there are two kinds of quantities: **scalar quantities,** or **scalars,** which have magnitude only, such as time, mass, and speed; and **vector quantities,** or **vectors,** which have both magnitude and direction, such as displacement, velocity, and acceleration.

Both scalars and vectors may be positive or negative with reference to some value taken (usually arbitrarily) as a zero value. The term **sense** is used as the geometrical equivalent of **sign,** and applies to both vectors and scalars. Thus one may travel along a road whose direction is north and south. Along this direction he may travel either northward (positive sense) or southward (negative sense). An electromotive force (Sec. 401), which is a scalar, may be so generated in a coil as to produce a current either counterclockwise (positive sense) or clockwise (negative sense) around the coil, but it has no relation to any particular direction in space.

A vector may be represented by a straight line whose length represents (to some convenient scale) the magnitude of the vector, whose direction is parallel to the direction of the vector, and whose arrow point indicates the sense of the vector.

The importance of the distinction between scalars and vectors will be seen from the following examples.

Scalars: A man having 4 lb of sugar (a) buys (addition of positive quantity) or (b) sells (addition of negative quantity) 3 lb of sugar. These transactions are represented by the equations:

$$\text{(a)}\ 4\text{ lb} + 3\text{ lb} = 7\text{ lb}$$
$$\text{(b)}\ 4\text{ lb} - 3\text{ lb} = 1\text{ lb}$$

and these two are the only ways in which the scalar quantities, 4 lb and 3 lb mass, can be added.

Vectors: A boat has a velocity of 4 ft/min eastward. A man on deck (a) walks 3 ft/min eastward, or (b) walks 3 ft/min westward, or (c) walks 3 ft/min northward.

These cases are represented as follows:

(a) 4 ft/min + 3 ft/min = 7 ft/min eastward

(b) 4 ft/min − 3 ft/min = 1 ft/min eastward

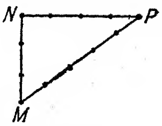


FIG. 19

(c) Since in one minute the man walks 3 ft (MN, Fig. 19) northward on the boat, while at the same time the boat carries him 4 ft NP eastward, at the end of 1 min he finds himself at P, a distance of 5 ft northeast of his initial position M, marked by an anchored buoy, say.

The addition of velocities of 4 ft/min and 3 ft/min gives velocities of 7 ft/min, 1 ft/min, or 5 ft/min, depending upon whether the given velocities are in the same direction (addition of positive quantity), in the opposite direction (addition of negative quantity), or at right angles.

21. Vector addition and subtraction. The above example shows that the processes of ordinary (scalar) algebra are not sufficient for problems involving vectors. We therefore devise a method of vector addition and subtraction. Consider a particle at P_0 (Fig. 20) and let it be given two displacements P_0M and P_0N respectively. If the displacement P_0M is given first and then MP (= and || to P_0N), the final position P will be exactly the same point as if P_0N is given first and then NP (= and || to P_0M).

In either case the result of the two displacements is that the particle is displaced from the point P_0 to the point P.

Therefore, by the definition of displacement (Sec. 17), P_0P is the displacement resulting from (i.e., the sum of) the two displacements, if given consecutively. From the above reasoning and from similar parallelograms, if the particle is given the displacement P_0M' and then $M'P'$, which are half of P_0M and P_0N, respectively, it will then occupy the position P', the midpoint of the diagonal P_0P, regardless of which is taken first.

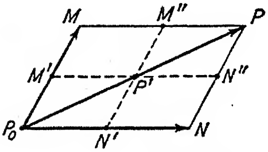


FIG. 20. Addition of Vectors

Obviously, if we take consecutive displacements which are the same fractions (say, 1/1,000,000) of P_0M and P_0N, after two such displacements the particle will be on the diagonal at a point whose distance from P_0 is 1/1,000,000 the distance P_0P; and a corresponding result will be true regardless of how small a fraction we take.

In the limit, by taking successively the same infinitesimal fractions of the given displacements, the particle would traverse the

diagonal. But this would appear to be equivalent to giving to the particle the displacements P_0M and P_0N simultaneously.

Hence we should expect the diagonal to represent the result, or sum, of the two given displacements, whether they are given consecutively or simultaneously. This is borne out by experiment.

While the foregoing discussion is with reference to displacements, it is clearly applicable to any vectors. Thus, we may state the law of parallelogram of vectors: **The sum, or resultant, of two vectors is represented by the concurrent diagonal of the parallelogram constructed on the two vectors as adjacent sides.**

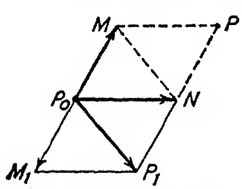


FIG. 21. Subtraction of Vectors

Subtraction being addition of a negative quantity, in Fig. 21 we would subtract vector P_0M from P_0N by adding P_0M_1, the negative of P_0M (i.e., equal in magnitude and opposite in direction), to P_0N. Hence P_0P_1 is the difference of the vectors P_0N and P_0M, for it is the vector sum of P_0N and the negative of P_0M. From the geometry of the figure it will be seen that P_0P_1 is equal and parallel to MN of the upper parallelogram. So, **the difference of two vectors is represented by the nonconcurrent diagonal of the parallelogram constructed on the two given vectors as adjacent sides.**

Graphical and trigonometric solutions. For many purposes vector sums or differences may be obtained satisfactorily by scaling the proper diagonals in figures accurately drawn to scale; but where greater accuracy is required, the solution is best made by the **law of cosines.**

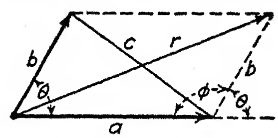


FIG. 22. Trigometric Method

In trigonometry (Fig. 22),

$$r^2 = a^2 + b^2 - 2ab\cos\phi$$
$$= a^2 + b^2 - 2ab\cos(180 - \theta)$$
$$= a^2 + b^2 + 2ab\cos\theta \qquad (14)$$

and
$$c^2 = a^2 + b^2 - 2ab\cos\theta. \qquad (15)$$

From these it is seen that the law of cosines gives the vector sum when the product term is added, and the vector difference when that term is subtracted.

Since the diagonal of a parallelogram divides it into two equal triangles, it is obviously necessary to draw only one of the trian-

gles. The parallelogram law then becomes the **law of triangle of vectors**: If two vectors v_1 and v_2 (Fig. 23) are represented to scale by two sides of a triangle, the second beginning at the end of the first, their vector sum, or resultant R, is represented to the same scale by the third side of the triangle, and its direction will be away from the origin O.

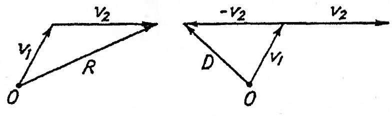


FIG. 23

The vector difference D is obtained by reversing the vector that is to be subtracted, and then proceeding as for a sum. Thus, in Fig. 23, $D = v_1 - v_2$.

Solved Problems

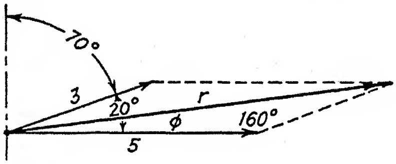


FIG. 24

1. If the wind gives a boat a velocity of 5 mph due east, and the current at the same time gives it a velocity of 3 mph N 70° E, what is its total velocity with reference to the shore (Fig. 24)?

Known:

$$\begin{aligned} a &= 5 \text{ mph} \\ b &= 3 \text{ mph} \\ \theta &= 20^\circ \\ r^2 &= a^2 + b^2 + 2ab\cos\theta. \end{aligned}$$

Solution:

$$\begin{aligned} r^2 &= 5^2 + 3^2 + 2 \times 5 \times 3 \times \cos 20^\circ \\ &= 25 + 9 + 30\,(.9397) \\ &= 62.19 \\ r &= 7.88 \text{ mph}. \end{aligned}$$

To find the direction of this resultant, we apply the law of sines to the figure.

$$\frac{\sin\phi}{\sin 160^\circ} = \frac{3}{7.88}$$

$$\sin\phi = \frac{3}{7.88} \times \sin 160^\circ$$

$$= \frac{3}{7.88} \times 0.3420 = 0.1302$$

$$\phi = 7^\circ 29'.$$

Hence the resultant velocity is 7.88 mph, N 82°31′ E.

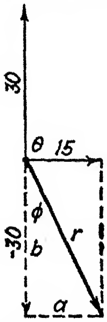


Fig. 25

2. An automobile is traveling with a velocity of 30 mph north. The wind blows east at 15 mph. Find the apparent velocity and direction of the wind relative to a person in the car (Fig. 25).

Analysis: If the car were at rest, a particle of air striking a passenger would appear to have the velocity 15 mph east, because of the wind alone. If there were no wind, the same particle of air would appear to have a velocity of 30 mph **south** because the car is rushing through the air at 30 mph **north.**

The resultant apparent velocity will be found by adding 15 mph east and 30 mph south, which is the same as subtracting 30 mph north.

Known:

$$a = 15 \text{ mph east}$$
$$b = 30 \text{ mph north}$$
$$\theta = 90°.$$

For vector subtraction,

$$r^2 = a^2 + b^2 - 2ab \cos \theta.$$

Solution:

$$\begin{aligned} r^2 &= 15^2 + 30^2 - 2 \times 15 \times 30 \times \cos 90° \\ &= 225 + 900 - 2 \times 15 \times 30 \times 0 \\ &= 1125 \\ r &= 33.5 \text{ mph.} \end{aligned}$$

To find the direction of the resultant, we have:

$$\tan \phi = \frac{a}{b} = \frac{15}{30} = 0.5$$

$$\phi = 26°34'.$$

Hence the velocity of the wind relative to a passenger is 33.5 mph, S 26°34′ E.

22. Polygon of vectors. When more than two vectors are to be added, the resultant of any two (v_1 and v_2) is found by the triangle as above. This partial resultant R_{12} is then combined with a third vector v_3 to give a second partial resultant R_{123}, and this process is continued until all the components have been used (see Fig. 26).

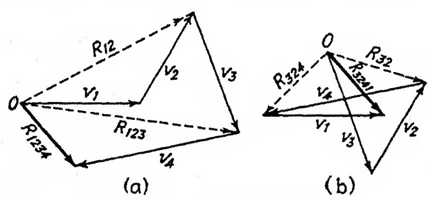


Fig. 26. Vector Polygons

It will now be observed that the resultant of all the vectors is the closing side of a polygon having the given component vectors

as sides, each component being drawn from the end of the preceding component, and the arrows pointing the same way around the polygon, except in the case of the resultant, which points oppositely to the others.

The law of the polygon of vectors may be stated: **If a number of vectors are drawn to scale to form the sides of a polygon, the arrows pointing in the same sense around the polygon, their resultant is represented in magnitude and direction by the closing side of the polygon, its arrow pointing away from the origin.**

The order in which the vectors are taken to form the polygon is quite immaterial, and the polygon may be crisscrossed by the vectors; but the final resultant will be the same in every case. Thus, Fig. 26*b* gives the resultant R_{3241}, which is exactly equal and parallel to R_{1234} as obtained in Fig. 26*a*, the subscripts indicating the order in which the vectors were used in each case.

As will be seen later, the polygon of vectors gives the resultant in magnitude and direction, but **it does not give the line of action unless the vectors are concurrent.** (See Sec. 66, Corollary.)

23. Vector representation of rotation. Since rotational displacements, velocities, and accelerations have magnitude, direction, and sense (sign), they are vectors. The conventional method of drawing the vector to represent one of these quantities is to take a line parallel to the axis of rotation whose length represents the magnitude according to some convenient scale; then place the arrow point on this line so that on looking along the axis in the direction of the arrow the rotation appears clockwise.

For example, in Fig. 18, if the angular displacement θ_1 is 60°, the line *MN* represents this vector to the scale: 1 cm = 20°.

24. Resolution of a vector into components. The **components** of a given vector are two or more vectors whose resultant would be the given vector.

A vector v (Fig. 27) may be resolved into two components v_1 and v_2 parallel to any two given lines *AB* and *CD* by constructing on v as the diagonal a parallelogram whose sides are parallel to *AB* and *CD*, respectively.

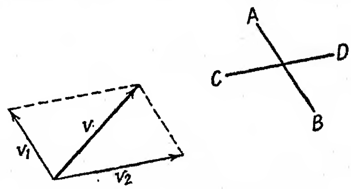


FIG. 27. Resolution of Vectors

25. Resolution of a vector into rectangular components. The most common case is that in which it is desired to resolve a vector v (Fig. 28) into two components v_x and v_y which are perpendicular to each other. For convenience, the usual axes of x and y are chosen, parallel to the two required directions. Then if v makes the angle θ with the x-axis, we have, from the definitions of sine and cosine:

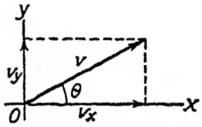


FIG. 28. Rectangular Components

$$v_x = v \cos \theta$$
$$v_y = v \sin \theta.$$

Solved Problem

A horse pulls on the traces of a sleigh with a force of 150 lb. If the traces make an angle of 10° with the horizontal road, what part of the force is effective in drawing the sleigh and what part merely lifts the forward end? (Fig. 29.)

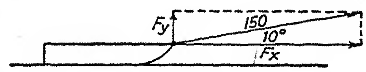


FIG. 29

Known:

$$\begin{aligned} F &= 150 \text{ lb force} \\ \theta &= 10° \\ F_x &= F \cos \theta \\ F_y &= F \sin \theta. \end{aligned}$$

Solution:

$$\begin{aligned} F_x &= 150 \times \cos 10° \\ &= 150 \times 0.985 \\ &= 147.8 \text{ lb force} \\ F_y &= 150 \times \sin 10° \\ &= 150 \times 0.174 \\ &= 26.1 \text{ lb force.} \end{aligned}$$

Hence, 147.8 lb force pulls the sleigh forward, and 26.1 lb force lifts the front of the sleigh.

26. Specifications of a motion. In general, the motion of a particle is completely determined when its path, its position on the path, its velocity, and its acceleration at any instant are known. The path is best described by its equation, as in analytic geometry, but often we merely state the kind of curve followed: e.g., a baseball thrown into the air describes a parabola. It will be noticed that these four necessary facts are found for each of the special types of motion discussed in the following paragraphs.

27. Rectilinear motion. Consider a particle moving in a straight line OX and having a **constant acceleration** a (Fig. 30). If its velocity is v_1 when the clock reads t_1, and v_2 when the clock reads t_2, we have:

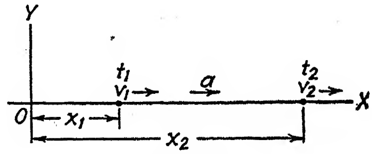


FIG. 30

$$a \equiv \frac{v_2 - v_1}{t_2 - t_1} \text{ by definition (Sec. 19),}$$

and calling the interval of time $(t_2 - t_1) \equiv t$ for brevity, we can write Eq. (10) thus:

$$a = \frac{v_2 - v_1}{t} \tag{16}$$

whence

$$v_2 = v_1 + at \tag{17}$$

which gives the final velocity v_2 in terms of the initial velocity v_1, the acceleration a, and the time t.

To find the position x_2 of the particle on its path at the clock time t_2, we revert to the definition of average speed, Eq. (5):

$$\bar{v} \equiv \frac{x_2 - x_1}{t_2 - t_1} = \frac{x_2 - x_1}{t}. \tag{18}$$

Clearing fractions,

$$x_2 - x_1 = \bar{v}t \tag{19}$$

$$x_2 = x_1 + \bar{v}t \tag{20}$$

which gives the final position in terms of the initial position x_1, the average velocity $\bar{v}$, and the time t.

But since the acceleration is constant,* the average velocity $(\bar{v})$ is given also by

$$\bar{v} = \frac{v_2 + v_1}{2} \quad \text{(definition of arithmetic mean).}$$

Therefore, from Eq. (19)

$$x_2 - x_1 = \frac{v_2 + v_1}{2}\,t.$$

Substituting for v_2 its value from Eq. (17),

$$x_2 - x_1 = \frac{(v_1 + at) + v_1}{2}\,t.$$

$$= \frac{2v_1 + at}{2}\,t$$

$$= v_1 t + \tfrac{1}{2}at^2 \tag{21}$$

* Otherwise the statement is not always true.

and

$$x_2 = x_1 + v_1 t + \tfrac{1}{2}at^2 \tag{22}$$

where

x_1 is the displacement when the time was t_1;
$v_1 t$ is the displacement due to the initial velocity v_1; and
$\frac{1}{2}at^2$ is the displacement due to the acceleration.

The acceleration a is + or − according as it is in the same direction as v_1 or opposite to it.

If we know along what straight line the particle moves, these formulas completely describe its motion, for we have:

1. The path is the straight line OX.
2. Its position at any instant t_2 is x_2, by Eq. (22).
3. Its velocity at any instant t_2 is v_2, by Eq. (17).
4. Its acceleration at any instant is constant ($= a$) by hypothesis.

It will be observed that the distance traversed in the time t is $x_2 - x_1$. If we call this distance s, i.e., $x_2 - x_1 \equiv s$, then from Eq. 19

$$s = \bar{v}t \tag{23}$$

and from Eq. 21,

$$s = v_1 t + \tfrac{1}{2}at^2. \tag{24}$$

28. Special case. When the particle starts from rest, $\mathbf{v}_1 = 0$; hence we may call $v_2 \equiv v$ without ambiguity, and Eqs. (17) and (24) take the much-used forms:

$$v = at \tag{25}$$

and

$$s = \tfrac{1}{2}at^2. \tag{26}$$

Corollary I. Equations (17), (20), and (22) are also correct when the a's, v's, and x's are vectors, not in a straight line, and the addition (or subtraction) is performed vectorially.

Corollary II. Another useful relation is obtained by eliminating t from Eqs. (16) and (23). This is most easily accomplished as follows:

From Eq. (23),

$$s = \bar{v}t = \left(\frac{v_2 + v_1}{2}\right)t$$

and from Eq. (16),

$$a = \frac{v_2 - v_1}{t}.$$

Multiplying, $$as = \frac{v_2 + v_1}{2}(v_2 - v_1)$$

whence $$v_2^2 - v_1^2 = 2as. \tag{27}$$

Solved Problems

1. With what speed v will a body reach the earth if it falls freely from rest at a height h above the earth?

Calling g the acceleration given to a mass by its own weight, we have, since the body starts from rest:

$$\begin{aligned} v_1 &= 0 \\ a &= g \\ s &= h \\ v_2 &= v. \end{aligned}$$

Substituting these values in Eq. (27),

$$\begin{aligned} v^2 - 0 &= 2gh \\ v &= \sqrt{2gh}. \end{aligned} \tag{28}$$

2. A body of mass M slides down a frictionless plane under the action of its weight alone. If the plane has the length s and makes an angle ϕ with the horizontal, what is the velocity of the body at the foot of the plane?

Analysis: Resolve the acceleration g which the body's weight would give it, if allowed to fall freely, into a component a_p parallel to the plane and a component a_n normal to the plane.

These components are:

$$\left.\begin{aligned} a_p &= g \sin \phi \\ a_n &= g \cos \phi \end{aligned}\right\} \text{from Sec. 25.}$$

Only the parallel component will result in motion.

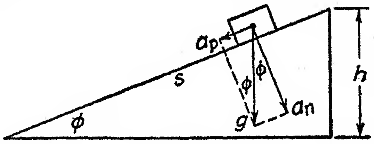


FIG. 31

Known:

$$\begin{aligned} v_1 &= 0, \text{ since the body starts from rest} \\ v_2 &\equiv v, \text{ which is to be found} \\ a &\equiv a_p = g \sin \phi \\ v_2^2 - v_1^2 &= 2as \text{ by Eq. (27).} \end{aligned}$$

Solution:

$$v^2 - 0 = 2(g \sin \phi)s = 2g\,(s \sin \phi).$$

From Fig. 31,

$$s \sin \phi = h$$

where h is the altitude of the plane.

Therefore,

$$v^2 = 2gh$$

and $$v = \sqrt{2gh}. \tag{29}$$

Comparing this with the preceding problem, it will be seen that the body reaches

the ground with the same speed whether it falls freely or slides down an incline without friction—or down any frictionless path.

3. A body has an initial velocity of 40 ft/sec and an acceleration of -2 ft/sec². How long before it will be 300 ft from the starting point? What will be its final velocity?

Known:

$s = 300$ ft

$v_1 = 40$ ft/sec

$a = -2$ ft/sec², $(-)$ meaning that the direction is opposite to that of the initial velocity

$$s = v_1t + \tfrac{1}{2}at^2. \qquad \text{by Eq. (24)}$$

Solution:

$$\begin{aligned} 300 &= 40t + \tfrac{1}{2}(-2)t^2 \\ t^2 - 40t &= -300 \\ t^2 - 40t + 400 &= 400 - 300 = 100 \\ (t-20)^2 &= 100 \\ t &= 20 \pm 10 \\ t &= 30 \text{ sec or } 10 \text{ sec.} \end{aligned}$$

For final velocity,

$$v_2 = v_1 + at. \qquad \text{by Eq. (17)}$$

When

$$\begin{aligned} t &= 30 \text{ sec,} \\ v_2 &= 40 + (-2)\,30 = 40 - 60 \\ &= -20 \text{ ft/sec opposite to the direction of initial velocity.} \end{aligned}$$

When

$$\begin{aligned} t &= 10 \text{ sec,} \\ v_2 &= 40 + (-2)10 = 40 - 20 \\ &= 20 \text{ ft/sec in the same direction as the initial velocity.} \end{aligned}$$

Hence the body will reach a point 300 ft from the starting point in 10 sec, moving in the direction of its initial velocity. It will then pass the point, finally reverse its motion, and again pass through the same point 30 sec after starting; but its velocity will then be opposite to the direction at first.

4. A body slides down a frictionless plane with an acceleration of 50 cm/sec². If the length of the plane is 3 meters, what must the initial velocity be in order that the body may reach the bottom with a velocity of 200 cm/sec?

Known:

$$\begin{aligned} a &= 50 \text{ cm/sec}^2 \\ s &= 3 \text{ m} = 300 \text{ cm} \\ v_2 &= 200 \text{ cm/sec} \\ v_2^2 - v_1^2 &= 2as. \end{aligned}$$

Solution:

$$\begin{aligned} (200)^2 - v_1^2 &= 2\,(50)(300) \\ v_1^2 &= 40{,}000 - 30{,}000 = 10{,}000 \\ v_1 &= \pm 100 \text{ cm/sec.} \end{aligned}$$

Thus the initial velocity may be 100 cm/sec down the plane, or 100 cm/sec up the plane (i.e., -100 cm/sec). In the former case it would reach the bottom after 2 sec; in the latter, after 6 sec—and in each case with the stipulated final velocity of 200 cm/sec. In the second case, it would go up the plane a certain distance, then reverse, and finally reach the bottom in 6 sec.

29. Curvilinear motion. The foregoing equations may be applied to any path, whether straight or not, if it is understood that a, v, and s are all measured along the path. In that case v would be called the **speed** and a would be the **tangential component only** of the particle's acceleration. The other (normal) component of the particle's acceleration is not accounted for by these equations. For that, see Sec. 33.

30. Rotation about a fixed axis. The analogue in rotation, of the motion discussed in the preceding paragraph, is rotation about a fixed axis with constant angular acceleration. For the direction of a rotation is defined (Sec. 23) as the **direction of the axis of rotation.** Hence, if the direction of the axis is fixed, the direction of the rotation is unchanged—although the sense (i.e., clockwise or counterclockwise) of the motion about this fixed axis may change.

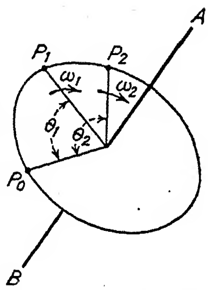


FIG. 32. Rotation about Fixed Axis

Consider a body rotating with constant angular acceleration α about an axis AB (Fig. 32), fixed in direction; and let the angular displacements and velocities be θ_1 and θ_2, ω_1 and ω_2, at clock times t_1 and t_2, respectively. Then,

$$\alpha \equiv \frac{\omega_2 - \omega_1}{t_2 - t_1} \qquad \text{from Eq. (11)}$$

and calling the time interval $(t_2 - t_1) \equiv t$ for brevity, as before, we have:

$$\alpha \equiv \frac{\omega_2 - \omega_1}{t} \tag{30}$$

whence

$$\omega_2 = \omega_1 + \alpha t \tag{31}$$

which gives the final angular velocity ω_2 in terms of the initial angular velocity ω_1, the acceleration α, and the time t.

It will now have been noticed that, in rotation, θ_1, θ_2, ω_1, ω_2, $\bar{\omega}$, and α correspond exactly to x_1, x_2, v_1, v_2, v, and a in translation.

Hence, by steps exactly analogous to those of Sec. 27, we deduce similar relations for rotation:

$$\theta_2 - \theta_1 = \bar{\omega}t; \qquad \bar{\omega} = \frac{\omega_2 + \omega_1}{2}$$

$$\theta_2 = \theta_1 + \omega t + \tfrac{1}{2}\alpha t^2.$$

Calling

$$\theta_2 - \theta_1 \equiv \theta, \qquad \theta = \bar{\omega}t$$

$$\theta = \omega t + \tfrac{1}{2}\alpha t^2, \tag{32}$$

and like Sec. 28, Corollary II,

$$\omega_2^2 - \omega_1^2 = 2\alpha\theta. \tag{33}$$

Solved Problem

What must be the angular acceleration of the rotor of a motor in order that in 5 sec it may come up to the speed of 1800 rpm from rest?

Known:

$$\begin{aligned}\omega_2 &= 1800 \text{ rpm} = 30 \text{ rps}\\ &= 2\pi(30) = 188.5 \text{ rdn/sec}\\ t &= 5 \text{ sec}\\ \omega_1 &= 0, \text{ since motor starts from rest}\\ \omega_2 &= \omega_1 + \alpha t.\end{aligned}$$

Solution:

$$188.5 = 0 + \alpha(5)$$

$$\alpha = \frac{188.5}{5} = 37.7 \text{ rdn/sec}^2.$$

31. Relations of angular and circumferential motions. At this point it is worth while to notice some important relations between the angular motion of a radius and the linear motion of a point on the radius.

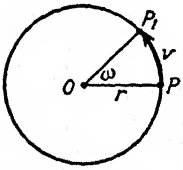


FIG. 33

When the circumferential speed v of a point P (Fig. 33) is constant, in 1 second P will describe an arc of length v while OP will sweep up the angle POP_1. But POP_1, being the angle turned through in 1 sec, is the angular speed ω of the radius; hence, by the definition of angle in circular measure;

$$\omega = \frac{v}{r} \tag{34}$$

and

$$v = r\omega. \tag{35}$$

Or, more simply, from Eq. (4),

$$s = r\theta.$$

Differentiating with respect to time,

$$v \equiv \frac{ds}{dt} = r\frac{d\theta}{dt} = r\omega.$$

Similarly,

$$a \equiv \frac{dv}{dt} = r\frac{d\omega}{dt} = r\alpha.$$

So we have:

$$\left.\begin{aligned} s &= r\theta \\ v &= r\omega \\ a &= r\alpha. \end{aligned}\right\} \qquad (36)$$

32. Direction of motion on a curve. If a particle is moving along a fixed curve, the direction of its motion at any instant must be the direction of the curve at the point where the particle is at that instant. But the direction of a curve at a point is defined in geometry as the direction of its tangent at that point. Hence the direction of the motion of the particle is the direction of the tangent to the curve at the then position of the particle (see Fig. 17). This definition is in accord with the fact that sparks from an emery wheel and particles of mud from an automobile wheel fly off along a tangent.

If the particle is to follow the curved path so as to arrive at P_1 (Fig. 34), it must obviously have an acceleration toward the curve (as a matter of fact, toward the center of curvature), for otherwise it would fly off along a tangent, or at least leave the curve. The magnitude and direction that this acceleration must have are deduced in the following section.

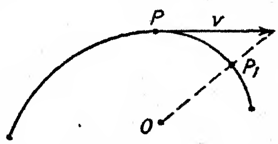


FIG. 34. Motion on a Curve

33. Uniform circular motion. Let a particle P (Fig. 35a) describe a circle about O as a center with constant linear speed v. Then, in t seconds, P will describe an arc P_1P_2 of length vt, and the radius OP will describe the angle $P_1OP_2 = \dfrac{vt}{r}$ radians, by Eq. (2).

Consequently, by Sec. 31, the angular speed ω of OP_1 is:

$$\omega = \frac{\left(\frac{vt}{r}\right)}{t} = \frac{vt}{rt} = \frac{v}{r} \text{ rdn/sec}$$

and it is constant since v and r are constant.

Make a velocity diagram (Fig. 35b) by drawing op_1 and op_2 equal and parallel to v_1 and v_2, respectively.* Then at represents the total change of velocity during the motion from P_1 to P_2 in magnitude and direction; for, by Sec. 21, the velocity at added vectorially to v_1 gives v_2.

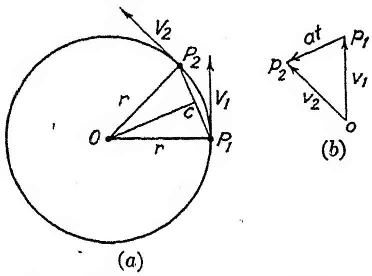


FIG. 35. Centripetal Acceleration

Then, since in magnitude $v_1 = v_2 = v =$ constant (by hypothesis), $\triangle\!\!\triangle\ P_1OP_2$ and p_1op_2 are similar.

(Two triangles are similar when an acute angle of one equals an acute angle of the other and the including sides are proportional.)

Therefore, for any value of t,

$$\frac{v}{r} = \frac{p_1p_2}{\text{chord } P_1P_2} = \frac{at}{\text{chord } P_1P_2}.$$

(Homologous sides of similar triangles are proportional.)

But the smaller t is taken, the more nearly is the chord P_1P_2 equal to the arc P_1P_2. Hence, in the limit as t approaches zero,

$$\frac{at}{\text{arc } P_1P_2} = \frac{v}{r}.$$

But arc $P_1P_2 = vt$. Therefore,

$$\frac{at}{vt} = \frac{v}{r}$$

$$a = \frac{v^2}{r} \tag{37}$$

or

$$a = \frac{v^2}{r^2}r = \omega^2 r. \tag{38}$$

* Drawing a separate velocity diagram is a mere matter of convenience to avoid confusion in Fig. 35a. It should be understood that facts found in Fig. 35b are equally true of Fig. 35a.

This acceleration is constant in magnitude, since ω and r are constant; but it is not constant in direction.

To find the direction of this acceleration, we notice from the geometry of the figures that p_1p_2 is always perpendicular to chord P_1P_2 for any value of t. Therefore, p_1p_2 is parallel to CO the perpendicular bisector of P_1P_2. But CO is always radial. Therefore p_1p_2, or at, being always parallel to CO, is also radial and is always directed toward the center. Since t is a scalar and has no direction, the direction of a is the same as that of at.

Hence **in uniform circular motion the acceleration is toward the center ("centripetal") and equals v^2/r.**

Since in this motion v is tangential, it is seen that the acceleration is always at right angles to the velocity.

As a check upon the nature of v^2/r let us see if its dimensional formula is the same as that of acceleration:

$$\frac{v^2}{r} = \frac{[LT^{-1}]^2}{[L]} = [LT^{-2}]$$

which is the same as that obtained for acceleration in Sec. 19.

The fact that two quantities have the same dimensional formula is a **necessary** but not a **sufficient** condition that they have the same nature. Work and torque (Sec. 54) both have the dimensional formula $[ML^2T^{-2}]$ but are distinctly different kinds of physical quantities.

Solved Problem

In the problem of Sec. 30, if a conductor of the rotor is 20 cm from the axis of rotation, what is its final linear speed and its total acceleration?

Known:

$$\left.\begin{aligned}\omega_2 &= 188.5 \text{ rdn/sec} \\ \alpha &= 37.7 \text{ rdn/sec}^2\end{aligned}\right\} \text{from solution in Sec. 30}$$

$$r = 20 \text{ cm}$$

$$v = r\omega \qquad \text{by Eq. (35)}$$

$$\text{Tangential component of } a = r\alpha \qquad \text{by Eq. (36)}$$

$$\text{Radial component of } a = \frac{v^2}{r}. \qquad \text{by Eq. (37)}$$

Solution:

$$v = 20 \times 188.5 = 3770 \text{ cm/sec}$$

$$\text{Tangential component of } a = 20 \times 37.7 = 754 \text{ cm/sec}^2$$

$$\text{Radial component of } a = \frac{(3770)^2}{20} = 710{,}645 \text{ cm/sec}^2$$

$$\text{Total acceleration } a = \sqrt{(754)^2 + (710{,}645)^2} = 710645.4 \text{ cm/sec}^2.$$

34. Parabolic motion. Another case in which the acceleration is not in the same direction as the velocity is the following.

As we shall see in Sec. 125, a body falling freely near the surface of the earth has nearly uniform acceleration toward the earth. Hence, if we consider a particle thrown out horizontally with a constant velocity v_0, and falling with a constant vertical acceleration a, we find at any time t after starting that the particle will be at some point P whose abscissa OA and ordinate OB are the distances traveled in virtue of its horizontal velocity v_0 and vertical acceleration a, respectively. (See Fig. 36.)

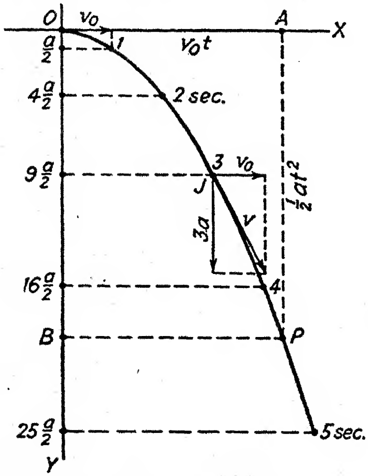


FIG. 36. Parabolic Motion

In the notation of analytic geometry,

$$x = v_0t \qquad \text{by Eq. (23)}$$

$$y = \tfrac{1}{2}at^2 \qquad \text{by Eq. (26)}$$

so that by eliminating t from these equations, we have the Cartesian equation of the path of the particle:

$$x^2 = \frac{2v_0^2y}{a} \tag{39}$$

which represents a parabola. The velocity v of the particle at any instant has the horizontal component, $v_x = v_0$, and the vertical component, $v_y = at$, by Eq. (17), which on being combined give v tangent to the parabola at J.

The acceleration is a and vertical, which is not in the direction of the motion for finite values of t.

In all the cases thus far considered the acceleration has been constant in magnitude, though not always constant in direction. An important case in which the acceleration is not constant in magnitude is discussed in the next section.

35. Simple harmonic motion. Motion in which the acceleration a is directly proportional to the displacement x, but in the opposite direction, is called simple harmonic motion. In nature, motions closely approximating this occur frequently. The motion

of the vibrating parts of musical instruments is of this kind, whence the term **simple harmonic motion** (shm).

By the definition,

$$a \propto -x.$$

where the minus sign means that a and x are in opposite directions.

Therefore, the algebraic expression of the definition is:

$$a = -cx \tag{40}$$

where c is a positive constant.

The meaning of the constant c and other relations in shm are most readily seen by the use of an **auxiliary circle** (Fig. 37).

Let a particle Q describe the auxiliary circle of radius r with constant speed V. The total acceleration a' of Q is from Q toward the center O and has the value (Sec. 33):

$$a' = \frac{v^2}{r} = \omega^2 r.$$

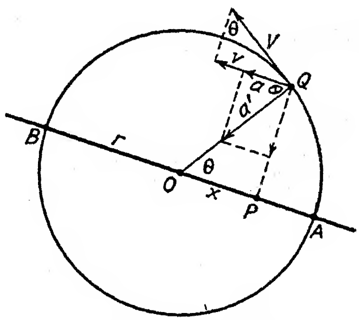


FIG. 37. Auxiliary Circle

Let P be the projection of Q on *any* straight line AB. The acceleration a of P will obviously be the same as the component of Q's acceleration parallel to AB, since P is always perpendicularly under Q.

Therefore,

$$a = -a' \cos \theta$$

the explanation of the negative sign being that, when the displacement x ($= OP$) of P is to the right, a is toward the left.

But

$$\cos \theta = \frac{x}{r}.$$

Therefore

$$a = -a' \frac{x}{r} = -\omega^2 r \left(\frac{x}{r}\right)$$

$$a = -\omega^2 x. \tag{41}$$

Hence the point P will have shm along the line AB, for its acceleration a conforms to the definition of shm in Eq. (40). Also, the meaning of the constant c has become clear: it is the square of the angular speed $(\omega \text{ rdn/sec})^2$ of the radius OQ of the auxiliary circle.

Since the acceleration of P is known, its motion at any time t sec after P leaves the end A of its travel will be completely known if we find v and x.

Resolving V parallel and perpendicular to AB, we have:

$v = -V \sin \theta$, since P must keep perpendicularly under Q,

and $x = r \cos \theta$, from the definition of cosine; since $OQ \equiv r$, which is called the **amplitude.**

But $\theta = \omega t$. by Eq. (32)

Hence the shm of P is completely described by the equations:

$$\left.\begin{aligned} a &= -\omega^2 x \\ v &= -V \sin \theta = -V \sin \omega t \\ x &= r \cos \theta = r \cos \omega t. \end{aligned}\right\} \tag{42}$$

It will be observed that P makes one round trip from A to B and back to A while Q makes one trip from A around the circle and back to A. Hence **the period T,** i.e., **the time of one complete vibration of P, is the same as the time it takes Q to make one trip around the auxiliary circle.**

The whole angle about the center O is 2π rdn, and since OQ describes ω rdn/sec, the time required for Q to go around is:

$$T = \frac{2\pi}{\omega}, \tag{43}$$

where T is also the **period** of the shm of the point P; and

$$\omega = \frac{V}{r}. \qquad \text{by Eq. (34)}$$

Further, from Eq. (42)

$$\omega = \sqrt{-\frac{a}{x}}$$

where the quantity under the radical is positive since a and x have opposite signs.

Therefore

$$T = 2\pi\sqrt{-\frac{x}{a}} = 2\pi\sqrt{-\frac{\text{Displacement}}{\text{Acceleration}}}. \tag{44}$$

Although these equations for shm have been derived for the case in which the path of the particle is a straight line, they are equally true for any path provided both x and a are measured along the path.

Solved Problem

An engine whose stroke is 20 in. runs at 600 rpm. Assuming that its crosshead has shm, what are its velocity and acceleration when 5 in. from one end of its stroke?

Known:

$$\begin{aligned} \omega &= 600 \text{ rpm} = 10 \text{ rev/sec} \\ &= 2\pi \times 10 = 62.83 \text{ rdn/sec} \\ r &= \tfrac{1}{2} \times 20 = 10 \text{ in.} \equiv \text{amplitude} \\ x &= 5 \text{ in.} \\ V &= r\omega \qquad\qquad \text{by Eq. (35)} \\ v &= -V\sin\theta \\ a &= -\omega^2 x. \qquad \text{by Eq. (42)} \end{aligned}$$

Solution:

In solving problems of shm it is usually best to begin by drawing the auxiliary circle and locating Q and P, as in Fig. 38.

$$\begin{aligned} \theta &= \cos^{-1}\tfrac{5}{10} = 60° \\ \sin\theta &= .866 \\ V &= 10 \times 62.83 = 628.3 \text{ in./sec} \\ v &= 628.3 \times .866 = -544.1 \text{ in./sec} \\ a &= -(62.83)^2 \times 5 = -19{,}740 \text{ in./sec}^2. \end{aligned}$$

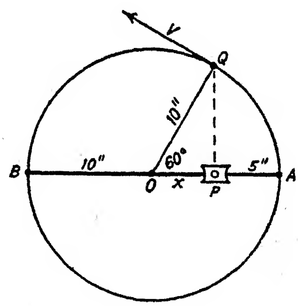


FIG. 38

36. Angular harmonic motion. Consider a rigid body vibrating back and forth about an axis perpendicular to the paper at A (Fig. 39).

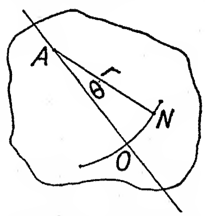


FIG. 39. Angular Harmonic Motion

Let its angular acceleration α at every instant be proportional to the angular displacement θ of any line AN from its initial position AO, and in the opposite sense. Then

$$\alpha \propto -\theta$$

$$\alpha = -c\theta. \tag{45}$$

If r is the radius of the arc on which *any* particle N oscillates, the linear acceleration a and linear distance s along this arc are:

$$a = r\alpha \qquad \text{by Eq. (36)}$$

$$s = r\theta. \qquad \text{by Eq. (2)}$$

Multiplying Eq. (45) by r,

$$r\alpha = -cr\theta$$
$$a = -cs$$

where s takes the place of x in Eq. (40).

But this is the defining equation of shm. Hence *every* particle of the body has simple harmonic motion on the circumference of a circle with center at A.

If we think of any of these circular paths as rectified, the numerical values of displacement, velocity, and acceleration of the particle along its path remain unchanged. An auxiliary circle could then be constructed as in Sec. 35, and we should find:

$$c = \omega^2.$$

Therefore

$$a = -\omega^2 s$$

and $$T = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{-\frac{a}{s}}} = 2\pi\sqrt{-\frac{s}{a}} = 2\pi\sqrt{-\frac{\theta}{\alpha}} \qquad (46)$$

for every particle, since c, or ω^2, is the same for all particles. Hence the equation:

$$\alpha = -c\theta \qquad (47)$$

is said to define **angular harmonic motion,** since every particle of a body having such motion will have shm as shown above.

PROBLEMS

1. A car covers a distance of 40 mi in 1.5 hr. Find its average speed in ft/sec.

2. A car leaving Richmond at 10:25 A.M. arrives in Norfolk at 1:15 P.M. If the distance is 125 mi, what is the average speed in mph and in ft/sec?

3. A car leaves New York at 3:45 P.M. and reaches Philadelphia at 6:12 P.M. If the distance is 98 mi, what is the average speed in mph and in ft/sec?

4. An automobile wheel is 30 in. in diameter and has a speed of 180 rpm. Find the speed of the car in mph.

5. The revolution counter of an engine indicated 168,000 revolutions at the end of an 8-hr run. What was the average speed in rpm and in rdn/sec?

6. At 9:45 A.M. the revolution counter of a winding engine read 28,314 and at 10.03 A.M. it read 36,432. Find the average speed in rpm and in rdn/sec.

7. If the earth describes its orbit (radius 93,000,000 mi) in 365 days, what is its average orbital speed in mi/sec?

8. The mileage-meter of a car is essentially a revolution counter. If the effective radius of the wheel is 12 in., how much should the meter reading increase when the wheel turns 5000 times? If this occurs in 6 minutes, what is the average reading of the speedometer?

9. If the diameter of the wheels in problem 8 is 2.1 ft, find the distance traveled and the average speed in mph.

10. At 8:11:21 the speed of a car is 15 mph and at 8:12:6 it is 45 mph. What is the average acceleration in mi/hr^2 and in ft/sec^2?

11. At 9:15:23 the speed of a car was 12 mph and at 9:17:11 its speed was 45 mph. What was the average acceleration in mi/hr^2 and in ft/sec^2?

12. A car going 50 mph is brought to a stop in 8 sec. What was the average deceleration?

13. A train having a speed of 60 mph is brought to rest in 11 sec. What is the average deceleration? What is the distance covered in that time?

14. A revolution counter on a bicycle read 26,116 at 8:10 A.M. and 27,244 at 8:18 A.M. Find the average angular speed in rpm and in rdn/sec.

15. An automobile engine speeds up from 120 rpm at 7:24:35 to 1200 rpm at 7:24:40. What is the average angular acceleration in rev/min^2 and in rdn/sec^2?

16. An electric motor speeds up from 150 rpm at 8:16:45 to 3600 rpm at 8:18:26. What is the average angular acceleration in rev/min^2 and in rdn/sec^2?

17. A wood shaper has a speed of 400 rpm at 8:45:15 and 10,000 rpm at 8:45:27. What is its average angular acceleration in rev/min^2 and in rdn/sec^2?

18. The shaft of a hoisting engine has on it two drums 18 in. and 24 in. in diameter, respectively. What is the speed of the hoisting ropes on the two drums when the shaft makes 120 rpm?

19. A speedboat changes its speed from 10 mph to 40 mph with a uniform acceleration of 20 mi/hr^2. What distance does it travel in that time?

20. A body moving with a speed of 15 mph is accelerated uniformly at the rate of 10 ft/sec^2 for 20 sec. Find the velocity at the end of that time and the distance covered in that time.

21. A motorboat changes its speed from 5 mph to 20 mph with a uniform acceleration of 6 mi/hr–min. How far does it go in that time?

22. An airplane changes its speed from 90 mph to 150 mph with a uniform acceleration of 30 mi/hr–min. How far does it go in that time?

23. The muzzle speed of a bullet from a rifle barrel 2 ft long is 1200 ft/sec. Find the acceleration, assuming that it is uniform, and the time to traverse the barrel. If the rifling makes 1.5 turns in the length of the barrel, what is the rotational speed of the bullet when it emerges?

24. An automobile engine will come from rest up to its rated speed of 1200 rpm in 1/4 min. What is its angular acceleration in rdn/sec^2? How many revolu-

tions does it make in that time? What is the linear speed of a point on the rim of the flywheel 8 in. from the center?

25. A sled having an initial velocity of 210 ft/sec is given an acceleration of -20 ft/sec². How long will it be before the sled is 540 ft from its initial position? Explain the two answers.

26. A car has an initial speed of 120 ft/sec. If it is given a uniform acceleration of -4 ft/sec², how long will it be before it is 550 ft from its initial position?

27. A ball is thrown vertically upward and is caught 4 sec later. How high did it rise? What was its velocity when caught? (Take $g = 32$ ft/sec².)

28. Under the action of gravity alone ($g = 980$ cm/sec²), a body slides from rest down a frictionless plane inclined 18° to the horizontal. What will be its speed at the end of 10 sec and how far does it go in the last 3 sec?

29. If two automobiles are moving in opposite directions with speeds of 40 and 50 mph, respectively, what is the speed of one relative to the other?

30. A ship is sailing due east at a speed of 20 knots when a passenger walks due north on its deck with a speed of 2 knots with respect to the ship. What is the velocity of the passenger relative to the earth?

31. From a train moving 60 mph eastward, a package is thrown with a velocity of 50 ft/sec N 45° E relative to the train. Find the resultant velocity of the package in magnitude and direction relative to the ground.

32. A river flows S 40° E at a speed of 4.5 mph. A man swims N 10° E at a speed of 2 mph relative to the water. What are the speed and the direction of the swimmer relative to the earth?

33. Two horses are hitched to a post so that the angle between their traces is 60°. If one horse pulls with a force of 300 lb and the other with a force of 500 lb, what is the force on the post?

34. If the traces to a wagon make an angle of 8° with the horizontal, how much force draws the wagon along and how much tends to lift it when the horse exerts a pull of 200 lb on the traces?

35. A boy pulls a sled along a level path by a rope which makes an angle of 20° with the horizontal. If he exerts a total force of 80 lb, how much force lifts the sled and how much pulls it forward?

36. A stream flows eastward with a velocity of 2.4 mph relative to the bank. A boatman rows S 30° W at 3.6 mph relative to the water. What is his resultant velocity relative to the bank?

37. A train is running north with a velocity of 45 mph and the wind is blowing west with a velocity of 15 mph. Find the velocity of the wind relative to the train.

38. An airplane pilot wishes to follow a course N 35° W. If the velocity of the wind is 30 mph N 42° W and the speed of the plane relative to the air is 150 mph, what should be the reading of his compass?

39. A billiard ball is given simultaneously the following velocities: 4 ft/sec southward; 3 ft/sec northeast; 6 ft/sec eastward; 5 ft/sec northwest. Find graphically the total (resultant) velocity by two different polygons.

40. A man rows a boat N 70° E at a speed of 3.2 mph with respect to the water of a river which flows due north at a speed of 2.7 mph relative to the bank. What is the velocity of the boat relative to the bank?

41. A push ball is given the following velocities simultaneously: 3 ft/sec northward; 4 ft/sec eastward; 7 ft/sec southeast; 6 ft/sec southwest. Find the total (resultant) velocity graphically by two different polygons.

42. Find the sum of three vectors of 3, 4, and 6 units making angles of 30°, 45°, and 60° respectively with the horizontal.

43. A soccer ball receives simultaneously the following velocities: 2 ft/sec southward; 3 ft/sec eastward; 4 ft/sec northward; 1 ft/sec westward. Find its total (resultant) velocity graphically by two different polygons.

44. If a hydrogen-driven turbine 6 mm in diameter rotates 21,600 rps, what is the linear speed and the acceleration of a particle on its periphery?

45. In the ultracentrifuge of Beams and Weed, a particle 0.4 in. from the axis has an acceleration 1,000,000 times that of gravity. What is the linear speed of the particle, and the angular speed of the rotor in rpm?

46. A rock is thrown horizontally with a speed of 50 ft/sec from the top of a cliff 256 ft high and just reaches the opposite bank of a river that runs at the foot of the cliff. What is the width of the river?

47. If an airplane which has a speed of 150 mph drops a bomb when the plane is at a height of 1280 ft and directly over the funnel of a ship at anchor, how far from the funnel will the bomb strike?

48. A river 210 ft wide runs at the base of a vertical cliff 144 ft high. With what horizontal velocity must a rock be thrown from the cliff in order to reach the opposite bank of the river?

49. The speed of a certain gas engine is 1200 rpm and its stroke is 8 in. Assuming the piston to have shm, what is its velocity and acceleration when 3 in. from one end of the stroke?

50. The stroke of a certain steam engine is 16 in. and its speed is 540 rpm. Assuming the motion of the crosshead to be shm, what is its velocity and acceleration when 4 in. from one end?

51. The blade of a reaper executes shm with an amplitude of 2 in. and a period of 0.1 sec. Find its speed and acceleration when 0.5 in. from the center of its stroke.

52. The needle of a sewing machine has approximately shm with a path length of 3 cm and makes 600 vib/min. What is its displacement and acceleration 1/30 sec after passing the center of its path?

53. A body whose mass is 4 lb is vibrating on a spring with a frequency of 1 vibration per second. What is its acceleration when it is 2 in. from the center of its travel?

DYNAMICS

37. Dynamics is that part of mechanics which deals with the action of force.* It is divided into **statics,** which treats of forces in equilibrium, and **kinetics,** which treats of forces not in equilibrium.

Force is defined as the time rate of change of momentum, or as the space rate of change of energy. These definitions mean the same thing and will become clear after Secs. 40 and 79 have been studied. When we say that a force acts upon a body we mean merely that the body is being accelerated.

For the present we may think of force as a **push** or a **pull.**

38. Methods of measuring force are many, but it will be sufficient here to mention four:

1. By means of the *spring balance.* If we draw a body B (Fig. 40) along by means of a spring balance S, thus causing it to

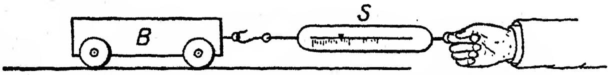


FIG. 40. Measuring Force with Spring Balance

speed up, we observe that the spring shows more elongation the more rapidly it causes the motion of the body to change. The spring balance is therefore a means of measuring force, and it is the one that will be used most frequently in this text because of its simplicity.

2. By the muscular effort required to do various things. This gives us our most primitive conception of the meaning of force, but is a very inexact method of measuring forces.

3. By the mass supported against the force of gravity. Thus if a rope or a post supports a mass of 1 lb, the rope or post exerts an upward force of 1 lb upon the body. This is a convenient method and satisfactory for many purposes.

* "Dynamics" is often used as a synonym for "mechanics."

4. By the acceleration given to a body. This is the basic method of measuring force, in fact of **defining** force, and was devised by Sir Isaac Newton. Before taking up the study of this method, three other physical quantities, linear momentum, linear impulse, and pressure, are defined.

(a) **The linear momentum of a body is defined as the product of its mass M by its velocity v.** This is for motion of translation,*

$$\text{Linear momentum} \equiv Mv. \tag{48}$$

Newton considered momentum to be the measure of quantity of motion. If we think of a car moving at the rate of 10 mph as having a certain amount of motion, then at 20 mph it might be said to have twice as much motion. Or, two cars moving at 10 mph might be said to have twice as much motion as one such car. Here momentum is clearly what we have in mind, so in a sense momentum is the quantity of motion.

(b) **Linear impulse is defined as the product of the force F by the time t that the force acts.** For motion of translation,

$$\text{Linear impulse} \equiv Ft. \tag{49}$$

Impulse is the measure of the time effect of a force.

Momentum and impulse are both vectors, for they are the products of the scalars, mass and time, by the vectors, velocity and force, respectively.

(c) **Pressure is defined as force per unit area.** Care should be taken not to use the term "pressure" when force is meant.

39. Newton's three laws of motion for translation.

1. A body at rest will remain at rest, and a body in motion will continue in motion in a straight line with constant velocity, unless acted upon by some external force.

2. Change of momentum is proportional to the impulse that produces it, and takes place in the direction of the force.

3. To every action there is an equal and opposite reaction.

These laws, published by Newton in his *Principia* (1687), are the foundation upon which the whole structure of physics is built. The first two were devised by Galileo; only the third is actually Newton's work.

* For rotation see Sec. 57.

40. Discussion of Newton's laws of motion.

First law. This law is sometimes called the "law of inertia," because it says in effect that all bodies have inertia. It is a qualitative law and states what happens to a body when *no* force acts upon it. Obviously, it cannot be verified experimentally.

Second law. This is a quantitative law and states what happens when *some* force acts. To put it in the more convenient algebraic form, let M be the mass of a body; v_1 its velocity at the clock time t_1; v_2 its velocity at the clock time t_2; and F the force acting during the time interval $(t_2 - t_1) \equiv t$. Then,

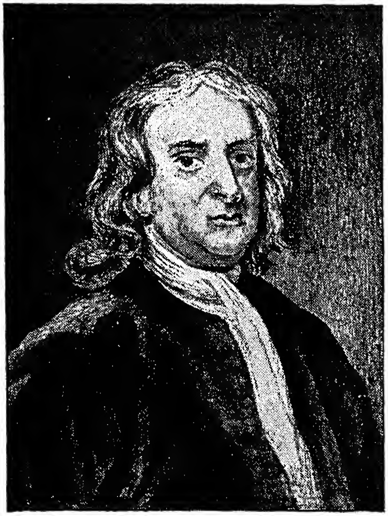


FIG. 41. Sir Isaac Newton

$$\begin{aligned}\text{Initial momentum} &= Mv_1\\ \text{Final momentum} &= Mv_2\\ \text{Change of momentum} &= (Mv_2 - Mv_1)\\ \text{Impulse} = F(t_2 - t_1) &= Ft.\end{aligned}$$

Hence, by Newton's second law,

$$Ft \propto (Mv_2 - Mv_1).$$

Therefore

$$Ft = k(Mv_2 - Mv_1)$$

and

$$F = kM\frac{(v_2 - v_1)}{t}. \tag{50}$$

But

$$\frac{v_2 - v_1}{t} \equiv a. \qquad \text{by Eq. (10)}$$

Consequently for any units whatever,

$$F = kMa \tag{51}$$

where the constant of proportionality k depends upon the choice of units for F, M, and a, as will be seen in Sec. 46; and these quantities are so defined as to make k a pure number. Equation (50) is in fact a scientific definition of force as it is stated in Sec. 37. The constant k affects the units only.

Third law. The words "action" and "reaction" are simply other names for **force**. Thus, this law states that for every force there is an equal and opposite force, or that forces always occur in equal opposed pairs. It should be carefully noted that **the two forces act upon different bodies**; i.e., A acts upon B and B reacts upon A. This is illustrated by the familiar fact that when a gun fires a bullet, the bullet reacts upon the gun, thus causing the "kick."

Again, when a body B is suspended from a support (Fig. 42), the earth acts downward on the body with the force of gravity W, which we call its weight, and the body reacts upward on the earth with the equal force W'. The body in turn acts downward on the rope with the force F and the rope reacts upward on the body with the equal force F'. F is clearly equal to W because the body is supported at but one point, and F' must equal F because the body does not fall.

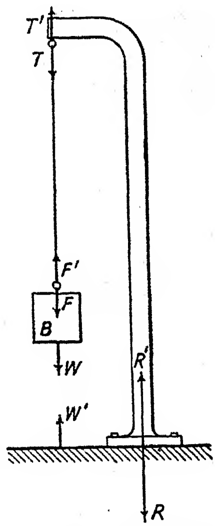


FIG. 42. Newton's Third Law—Body at Rest

Even when the bodies move, there is always this same equality of action and reaction, as is easily seen if a spring balance S is introduced into the line as shown below.

Let a sled M whose mass is 64 lb be attached to an automobile (Fig. 43), and let the friction f' of the ground on the sled be 36 lb force. As the rope is tightened, the action F of the spring balance on the sled and the reaction F' of the sled on the balance increase until the pull on the sled is just equal to the friction which opposes the motion; the spring balance then registers 36 lb. The reaction F' must equal F as otherwise the spring would not stay stretched.

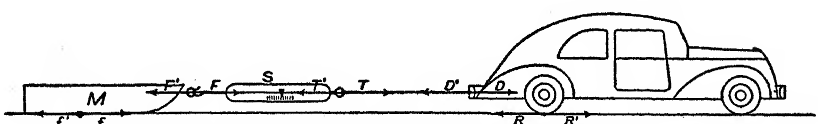


FIG. 43. Newton's Third Law—Body in Motion

If the tension T in the rope is then increased to, say, 40 lb force, the spring will stretch until the pull F on the sled is 40 lb and the

sled will speed up because the pull now exceeds the friction by 4 lb force. This excess of 4 lb force will give the sled an acceleration of 2 ft/sec^2, and the spring balance will continue to register 40 lb as long as this acceleration is maintained.

But the balance would not stay extended to 40 lb unless the reaction F' were 40 lb. Hence, even when the bodies are in motion, action and reaction are equal and opposite.

In this case the reaction F' equals the resistance due to friction f' and the resistance due to the inertia of the body, the latter being 4 lb force. If the car ceases to accelerate and maintains a uniform speed, the spring balance contracts until it reads 36 lb, just enough to overcome friction.

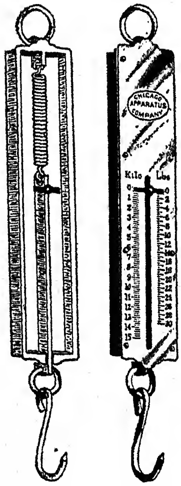


FIG. 44. Spring Balance

41. Weight, the force of gravity. One of the most familiar phenomena in nature is that unsupported bodies fall toward the earth and that their speed increases as they fall; i.e., they experience an acceleration toward the earth. We attribute this acceleration to the pull, or force of attraction, of the earth for the bodies.

The pull of the earth on a body is called the force of gravity, or weight,* of the body. Being a force, it may be measured with a spring balance (Fig. 44).

42. The direction of weight is obviously that of a plumb line, or vertical, hence the definitions:

Vertical. A vertical direction, or a vertical, is defined as the direction of a plumb line, i.e., a flexible thread suspending a heavy body (the bob). This direction is approximately toward the center of the earth, except in the neighborhood of a mountain or a large deposit of exceptionally dense material.

Horizontal. A line or a plane is said to be horizontal when it is perpendicular to a vertical. In Sec. 96 it will be shown that the surface of a liquid at rest is horizontal.

43. Inertia is that property exhibited by all bodies by which they resist any change in their motion, either in magnitude or in direction.

* The word "weight" as used in everyday business means the same as "mass" in science.

So fundamental in our notion of matter is this property of inertia that it is used to define matter: **Matter** is anything that has inertia.

Experiment shows that the passive resistance which a piece of matter offers to any change in its condition of rest or motion is proportional to its **mass** as obtained by weighing on a beam balance (see Fig. 3). Mass is therefore a measure of a body's inertia. Newton considered mass to represent the "quantity of matter" in a body.

The law of conservation of matter (or mass), first enunciated by Lavoisier in 1774, states that matter can be neither created nor destroyed. It appears well established, however, that matter may be converted into energy, and vice versa. In chemical reactions, Manley has found the law correct to 1 part in 100,000,000.

44. Contrasting mass and weight. Consider a piece of brass B (Fig. 45) whose mass is 1 lb as determined by a beam balance. We think of this pound of mass as the quantity of brass in the body.

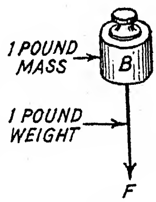


FIG. 45. Mass and Weight Distinguished

If the body is free it will fall toward the earth with acceleration. We say, therefore, that it is acted upon by a force, the force of gravity, or **weight**, of the body; and we may measure the weight with a spring balance.

If we should take the body, the beam balance, the set of standard masses, and the spring balance from the equator to the pole, and measure the mass and weight at many places along the way, we should find that the mass measurements would all be the same. However, the weight as shown by the spring balance would increase as we go from the equator toward the pole. The weight of a body is 1/289 less at the equator than it is at the pole. Hence,

The mass of a body refers to the amount of matter in it and is everywhere the same. The weight means the pull of the earth on the body and varies from place to place.

45. Gravitational units of force. In the British gravitational system, the unit of force is the pound of force: **A pound of force is equal to the weight of a pound of mass.**

Similarly, in the metric gravitational system, **a gram of force is equal to the weight of a gram of mass.**

These units are called gravitational units of force because they depend upon the force of gravity, or weight. But they do not have fixed values, for the weight of a unit of mass varies from place to place. There is, however, one advantage in using these gravitational units of force: the weight of the body is then always **numerically** equal to its mass. For example, if we have a body of 5 lb mass, obviously its weight will be 5 lb of force, for the pull of the earth on each pound of mass is defined as a pound of force.

Here will be seen the necessity of distinguishing carefully between a pound of mass (lbm) and a pound of force (lbf), since the word "pound" is used in *two* senses.

46. The constant k of Newton's second law. We are now able to determine the constant of proportionality k in Newton's second law,

$$F = kMa. \qquad \text{from Eq. (51)}$$

Consider the case of a freely falling body at Washington, D.C. From the table in Sec. 125 its acceleration is found to be 32.16 ft/sec². If the body falling is 1 lb of mass, we have:

$$\begin{aligned} M &= 1 \text{ lb mass} \\ F &= 1 \text{ lb force } (\equiv \text{weight of body}) \\ a &= 32.16 \text{ ft/sec}^2. \end{aligned}$$

Since Newton's second law is true for any case of a force giving acceleration to mass, it is true for this case. Substituting these values in Eq. (51),

$$1 = k \times 1 \times 32.16$$

$$k = \frac{1}{32.16}.$$

Hence, for the British gravitational system of units, k is the pure number 1/32.16 at Washington, D.C.

Similarly, when a gram of mass (gmm) falls freely it is acted upon by 1 gm of force (gmf), which gives it an acceleration of 980.1 cm/sec² (= 32.16 ft/sec²) at Washington, D.C. And applying Newton's second law as above, we find that, for the metric gravitational system, at Washington, D.C.

$$k = \frac{1}{980.1}.$$

From the foregoing it will be clear that the value of k depends upon the **choice** of units, and that **in both gravitational systems**

$$k = \frac{1}{g}$$

where g is the **pure number** numerically equal to the acceleration of gravity at the place where the experiment is made. Thus Newton's second law becomes, provided we use gravitational units,

$$F = \frac{1}{g} Ma. \tag{52}$$

That is, F is in lb force when M is in lb mass and a is in ft/sec²; and F is in gm force when M is in gm mass and a is in cm/sec².

47. Absolute units of force. The fact that the gravitational units of force are different at different places and require different values of k is a serious inconvenience in scientific work. To avoid these disadvantages, we devise a better system of force units which are everywhere the same and are, therefore, called **absolute units.**

Since k depends upon the **choice** of units, we can so choose our units as to make k anything we please. The simplest value it could have would be unity (= 1). Consequently we so choose these units as to make $k = 1$. In the British system the absolute unit of force is called the **poundal;** in the metric system, the **dyne.**

The $\left\{\begin{matrix}\text{poundal}\\ \text{dyne}\end{matrix}\right\}$ is defined as the force which acting alone upon one $\left\{\begin{matrix}\text{pound}\\ \text{gram}\end{matrix}\right\}$ of mass will give it an acceleration of one $\left\{\begin{matrix}\text{foot}\\ \text{centimeter}\end{matrix}\right\}$ per second per second.

We may visualize an imaginary experiment in which the conditions of these definitions are fulfilled. Let the body whose mass is 1 lb or 1 gm be upon a frictionless horizontal plane to which the force is parallel (Fig. 46) and in a perfect vacuum. The plane, being frictionless, offers no resistance to the motion; being perpendicular to the weight (which is vertical), there is no component of the weight parallel to the plane either to aid or to hinder the motion; and the experiment being in a perfect vacuum, there

is no air resistance. In these circumstances, the force being defined is **employed entirely in giving acceleration to the body.**

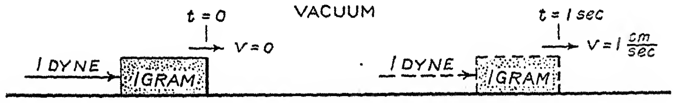


FIG. 46. Illustrating a Dyne

Let us now determine k for one of our absolute systems. From the definition of the dyne, we have the following data:

$$M = 1 \text{ gram mass}$$
$$F = 1 \text{ dyne}$$
$$a = 1 \text{ cm/sec}^2.$$

Substituting in Newton's second law, which is true for all cases,

$$1 = k \times 1 \times 1$$

whence

$$k = 1.$$

Obviously we shall get $k = 1$ in the absolute British system also.

Therefore we may write Newton's second law in the following simplified form, provided we use absolute units:

$$F = Ma \tag{53}$$

where F is in dynes when M is in gm mass and a is in cm/sec^2; and F is in poundals when M is in lb mass and a is in ft/sec^2.

Absolute units will now be seen to have *two advantages:* they are everywhere the same; they make the constant k unity in Newton's second law.

Dimensional formula for force. From Sec. 19, the dimensional formula for acceleration is $[LT^{-2}]$. Hence the dimensional formula for force is

$$[F] = [MLT^{-2}]. \tag{53a}$$

48. Relation of absolute to gravitational units of force. Let two forces F_1 and F_2 act separately on the same mass M and produce the accelerations a_1 and a_2, respectively. Then, by Eq. (53),

$$F_1 = Ma_1 \tag{a}$$

and

$$F_2 = Ma_2. \tag{b}$$

Dividing (a) by (b),

$$\frac{F_1}{F_2} = \frac{a_1}{a_2}. \tag{54}$$

Equation (54) states that *forces are proportional to the accelerations that they give to the same mass.*

We can apply this fact at once to determine the relation between the gram of force and the dyne. By their definitions we know that

1 gm of force gives 1 gm of mass an acceleration of 980.1 cm/sec² at Washington, D.C.

1 dyne gives 1 gm of mass an acceleration of 1 cm/sec² anywhere.

Hence, by Eq. 54,

$$\frac{1 \text{ gm force}}{1 \text{ dyne}} = \frac{980.1 \frac{\cancel{\text{cm}}}{\cancel{\text{sec}}^2}}{1 \frac{\cancel{\text{cm}}}{\cancel{\text{sec}}^2}}.$$

Therefore,

1 gm of force = 980.1 dynes at Washington, D.C.

Similarly,

1 lb of force = 32.16 poundals at Washington, D.C.

49. Weight in gravitational units. Since, in the gravitational systems of units, the unit of force is defined as the weight (at that place) of a unit of mass, it follows that for every unit of mass of a body there is a unit of weight. Therefore, its weight W is numerically equal to its mass M in these systems. For example, if a body has a mass of 10 gm, its weight is 10 gm of force: the two are *numerically* equal. That is,

$$\textbf{in gravitational units, } W = M \text{ (numerically only).} \tag{55}$$

50. Weight in absolute units. When a body falls freely, its weight W alone is the force that gives to its mass M the acceleration of gravity g. Since Newton's second law is applicable to all cases of a force acting alone upon a mass, we may put these values

in that law, Eq. (53). In this case, $F = W$; $M = M$; and $a = g$; so that,

$$\text{in absolute units,} \quad W = Mg. \tag{56}$$

Consequently, if the above body whose mass is 10 gm is at a place (e.g., Washington, D.C.) where the acceleration of gravity is 980.1 cm/sec², its weight expressed in absolute units is:

$$W = 10 \times 980.1 = 9801 \text{ dynes.}$$

Similarly, if a body has a mass of 10 lb, its weight in gravitational units is 10 lb of force according to Eq. (55); but in absolute units its weight at a place where the acceleration of gravity is 32.16 ft/sec² is

$$W = 10 \times 32.16 = 321.6 \text{ poundals.}$$

51. British engineering units. Since, when gravitational units are used, $k = 1/g$ and the weight W in pounds of force is **numerically** equal to the mass M in pounds of mass, engineers are accustomed to put these values into Eq. (51), which gives:

$$F = \frac{1}{g} Wa = \frac{W}{g} a. \tag{57}$$

Here W is force and g is the acceleration of gravity, so that W/g has the dimensions of mass and is called **the mass of the body in slugs.** That is, it defines a new unit of mass, the **slug** (or gee-pound), such that

$$1 \text{ slug} \equiv g \text{ pounds of mass.}$$

Using this gravitational unit of mass, which is **variable,** we may write as Newton's second law:

$$F = M''a \tag{58}$$

where F is the force in pounds;
M'' is the mass in slugs; and
a is the acceleration in ft/sec².

It will be seen that Eqs. (57) and (58) are exactly equivalent to Eq. (52).

Solved Problem

On an Atwood's machine (Fig. 47) two masses of 1 lb each are suspended by a massless cord over a frictionless and massless pulley. A mass of 0.1 lb is then added on one side. What is the resulting acceleration of the system, and what is the tension in the cord?

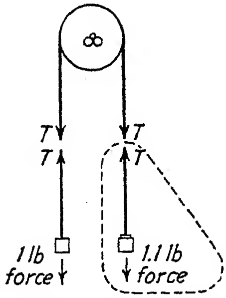


FIG. 47. Atwood's Machine

FIRST SOLUTION (*using absolute units*)

Known:

Total mass $M = 1 + 1 + 0.1 = 2.1$ lb mass
Unbalanced force $F = 0.1$ lb force $= 3.22$ poundals
$F = Ma.$ from Eq. (53)

Solution:

$$3.22 = 2.1a$$

$$a = \frac{3.22}{2.1} = 1.53 \text{ ft/sec.}^2$$

To find the tension T in the cord, imagine it cut and the tension applied on each side of the cut (Fig. 47). Then, considering the right-hand part in the dotted boundary, we have the following elements.

Known:

Total mass $M_1 = 1.1$ lb mass
Force upward $= T$ lb force
Force downward $= 1.1$ lb force
Unbalanced force $F_1 = (1.1 - T)$ lb force downward
$= 32.2\ (1.1 - T)$ poundals
$a = 1.53$ ft/sec^2 from above
$F_1 = M_1a.$

Solution:

$$32.2\ (1.1 - T) = 1.1 \times 1.53$$
$$T = 1.05 \text{ lb force.}$$

This value of the tension can be checked by cutting the cord on the left side, when

Total mass $M_2 = 1$ lb mass
Force upward $= T$ lb force
Force downward $= 1$ lb force
Unbalanced force $F_2 = (T - 1)$ lb upward
$= 32.2\ (T - 1)$ poundals
$a = 1.53$ ft/sec^2 from above solution
$F_2 = M_2a.$

Solution:

$$32.2\ (T - 1) = 1 \times 1.53$$
$$T = 1.05 \text{ lb force.}$$

SECOND SOLUTION (*using British engineering units*)

Known:

$$\text{Total mass } M'' = \frac{2.1}{32.2} \text{ slug} \qquad \text{(since } g = 32.2 \text{ ft/sec}^2 \text{ at Washington, D.C.)}$$

$$\text{Unbalanced force } F = 0.1 \text{ lb force}$$

$$F = M''a. \qquad \text{from Eq. (58)}$$

Solution:

$$0.1 = \frac{2.1}{32.2} \times a$$

$$a = \frac{3.22}{2.1} = 1.53 \text{ ft/sec.}^2$$

To find the tension T in the cord, we proceed as before to consider the right-hand part in the dotted boundary (Fig. 47).

Known:

$$\text{Total mass } M_1 = \frac{1.1}{32.2} \text{ slug}$$

$$\begin{aligned} \text{Force upward} &= T \text{ lb force} \\ \text{Force downward} &= 1.1 \text{ lb force} \\ \text{Unbalanced force } F_1 &= (1.1 - T) \text{ lb force downward} \\ a &= 1.53 \text{ ft/sec from above solution} \\ F_1 &= M_1''a. \end{aligned}$$

Solution:

$$(1.1 - T) = \frac{1.1}{32.2} \times 1.53$$

$$T = 1.05 \text{ lb force.}$$

These results are the same as were obtained in the first solution in which absolute units were used.

As an exercise, the student should check this value of T by carrying out the solution for the left side of Fig. 47, using B. E. (British engineering) units.

52. Conservation of momentum. Newton's laws of motion lead at once to one of the most important generalizations in physics: the law of the conservation of momentum.

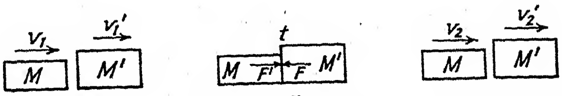


FIG. 48. Impact

Consider two masses M and M' (Fig. 48) moving with velocities v_1 and v_1' just before they collide, and with velocities v_2 and v_2' just after the collision. Let t be the time they are in contact.

During the contact period t, M will exert a force F' upon M'; and M', a force F upon M.

Taking directions to the right to be positive, we then have:

$$F = -F' \qquad \text{by Newton's third law;}$$

whence, by Newton's second law, Eq. (50):

$$kM\frac{(v_2 - v_1)}{t} = -kM'\frac{(v_2' - v_1')}{t}$$

$$Mv_2 - Mv_1 = -M'v_2' + M'v_1'$$

$$Mv_2 + M'v_2' = Mv_1 + M'v_1'. \tag{59}$$

Here the left side of the equation represents the total momentum of the system (consisting of the two bodies) after impact, and the right side represents the total momentum of the system before impact. Equation (59) therefore represents in mathematical symbols the **law of the conservation of momentum,** which may be stated in words as follows:

The total momentum of a system of bodies remains constant, provided that no force external to the system acts upon it.

Since momentum is a vector quantity, Eq. (59) means that if the two bodies are not moving along the same straight line, the **vector sum** of their momenta before impact equals the **vector sum** of their momenta after impact.

53. Line of action, lever arm. The straight line along which a force acts is called its **line of action.** From Fig. 49a it will be clear

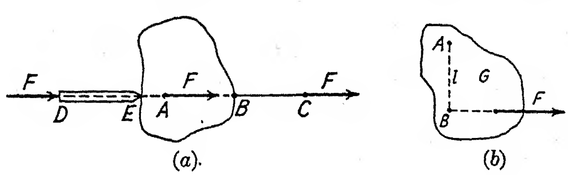


FIG. 49.

that **a force may be considered to act at any point in its line of action,** and its effect will be unaltered.

For example, we can tie a wire to the body at B and let the force F pull at the other end C of the wire; or we can use a strut DE and let the force push at the point D. The wire and the

strut, both of which must be considered weightless, transmit the force back to the body proper, and the effect is the same as if the force acted at A.

In Fig. 49b, let a body G be acted upon by a force F and so pivoted that it may turn about an axis perpendicular to the paper at A. Then the **lever arm**, or moment arm, of the force is defined as the perpendicular distance l from the axis to the line of action of the force.

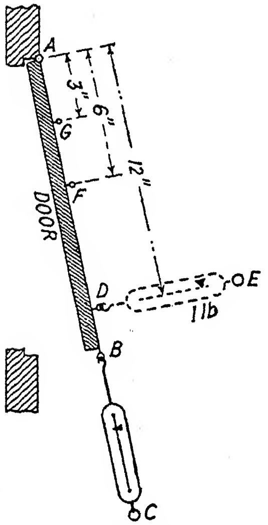


Fig. 50

54. Torque. The torque, or moment of a force, is its ability to produce rotation. In order to produce rotation, something more than mere force is required.

Consider a door AB (Fig. 50). If we attach a spring balance at B and pull so that the line of action of the force passes through the hinge A, no amount of force will cause the door to turn. But if, when we pull perpendicularly to the door at D, we find that it takes a force of 1 lb to turn the door uniformly, then when we pull similarly at F and at G, which are 6 in. and 3 in., respectively, from A, we shall find that forces of 2 lb and 4 lb, respectively, are necessary to turn the door uniformly.

Tabulating these observations:

	Force		Distance		Product
At D	1 lb	$\times$	12 in.	=	12 lb-in.
At F	2 lb	$\times$	6 in.	=	12 lb-in.
At G	4 lb	$\times$	3 in.	=	12 lb-in.

Here it is seen that in each case we get the same effect, namely, rotation of the door without acceleration, and the same product (12 lb-in.) of force by lever arm. We therefore infer that it is the product of a force by its lever arm that determines its rotational effect, and this is borne out by the most careful experiments.

Hence **torque $\mathfrak{T}$ is measured by the product of the force F and its lever arm l:**

$$\mathfrak{T} \equiv Fl. \tag{60}$$

55. Specifications of a force. It will be seen from Fig. 50 that the effect of a force upon a body depends upon

1. Its magnitude
2. Its direction
3. The location of its line of action.

In other words, the closer the force is to A, the less is its effect in producing rotation.

56. A couple is a system consisting of two forces equal in magnitude, parallel in direction, and opposite in sense (Fig. 51).

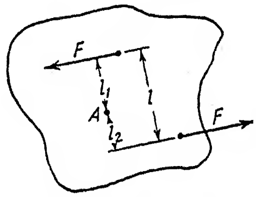


FIG. 51. A Couple

The **lever arm,** or moment arm, of a couple is the perpendicular distance l between the forces.

Taking moments about an axis perpendicular to the paper at any point A, we find that the torque $\mathfrak{T}$ of a couple is:

$$
\begin{aligned}
\mathfrak{T} &= Fl_1 + Fl_2 \\
&= F(l_1 + l_2) \\
&= Fl. \qquad (61)
\end{aligned}
$$

Hence the torque, or moment, of a couple equals the product of either force by the perpendicular distance between the forces.

The effect of a couple is to produce pure rotation. The effect is not altered by translation to any position in its plane or in any parallel plane; by rotation in its plane; or by changing the magnitude of its forces and the distance between them, provided the product of either force and the distance between them remains unaltered.

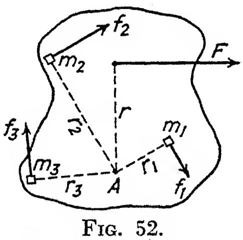


FIG. 52.
Moment of Inertia

A couple can be equilibrated only by another couple, equal in magnitude and opposite in direction to the given couple.

57. Newton's laws extended to rotation. Consider a rigid body (Fig. 52) free to turn about an axis perpendicular to the paper at any point A. Let the body be acted upon by a force F with lever arm r.

Imagine the force F replaced by little components $f_1, f_2, \ldots f_n$, acting upon the particles composing the body, whose masses are

$m_1, m_2, \ldots m_n$, respectively. It is assumed that the f's are so chosen that they will give the body exactly the same motion as was given it by F.

Then by Eq. (60)

$$\begin{aligned} \mathfrak{J} &= Fr \\ &= f_1r_1 + f_2r_2 + \cdots + f_nr_n \end{aligned} \tag{a}$$

in accordance with the above assumption.

By Newton's second law for translation (Eq. 53), we find that

$$\left.\begin{aligned} f_1 &= m_1a_1 \\ f_2 &= m_2a_2 \\ &\vdots \\ f_n &= m_na_n \end{aligned}\right\} \quad \text{(provided absolute units are used).}$$

Substituting these values of the f's in Eq. (a), above, we have:

$$\mathfrak{J} = m_1a_1r_1 + m_2a_2r_2 + \cdots + m_na_nr_n.$$

Since the body is rigid, all its parts will have the same angular acceleration α.

But

$$a_1 = r_1\alpha, \quad a_2 = r_2\alpha, \ldots a_n = r_n\alpha \qquad \text{by Eq. (36).}$$

Therefore,

$$\begin{aligned} \mathfrak{J} &= m_1r_1^2\alpha + m_2r_2^2\alpha + \cdots + m_nr_n^2\alpha \\ &= (m_1r_1^2 + m_2r_2^2 + \cdots + m_nr_n^2)\alpha. \end{aligned} \tag{62}$$

For brevity, we say that

$$(m_1r_1^2 + m_2r_2^2 + \cdots + m_nr_n^2) \equiv \Sigma mr^2. \tag{63}$$

The symbol Σ (Greek, sigma) means "the algebraic sum of." Σmr^2 therefore means the algebraic sum of the products obtained by multiplying each elementary mass by the square of its distance from the assumed axis. This quantity (Σmr^2) arises so frequently in mechanics that, for further brevity, it is given the symbol I and is called the "moment of inertia," or the "rotational inertia," of the body.

Moment of inertia is defined as the algebraic sum of the products obtained by multiplying each elementary mass by the square of its distance from the axis of rotation. In symbols,

$$I \equiv \Sigma mr^2. \tag{64}$$

Qualitatively, moment of inertia is a property of a body, depending upon its mass and shape, which determines the angular acceleration that will be given to the body when acted upon by a torque. In general, **the moments of inertia of a body about different axes will be different numbers.** (See table, Sec. 58.)

Using these symbols, Eq. (62) becomes

$$\mathfrak{T} = \alpha \Sigma m r^2$$

or

$$\mathfrak{T} = I\alpha \tag{65}$$

in which absolute units must be used because it has been derived from $F = Ma$.

If we carry out the above analysis using gravitational units, the constant k will come through the argument and we obtain the formula:

$$\mathfrak{T} = kI\alpha \tag{66}$$

which is analogous to Eq. (51).

Comparing Eq. (65) with Eq. (53), $F = Ma$, it will be seen that torque, angular acceleration, and moment of inertia in rotation correspond respectively to force, linear acceleration, and mass in translation. This correspondence will be found to continue, so that we call Eq. (65) or (66) **Newton's Second Law of Rotation,** which may be stated in words as follows: The time rate of change of the angular velocity of a body is proportional to the torque that produces it and takes place in the direction of the torque.

If $\mathfrak{T} = 0$ in Eq. (65) or (66), then α must be zero; hence we have **Newton's First Law of Rotation:** A body in rotation will continue in rotation with constant angular speed about an axis fixed in direction, unless acted upon by an external torque.

Continuing the analogy, we arrive at **Newton's Third Law of Rotation:** To every torque there is an equal and opposite reactive torque.

58. Moments of inertia about parallel axes (Lagrange's theorem). While it is true that moment of inertia in rotation corresponds to mass in translation, there is this important difference: **The mass of a body is always the same, but its moment of inertia depends upon the location of the axis of rotation.**

Often we know the moment of inertia of a body about a certain

axis, but need to know it about some other axis parallel to the first axis. In such a case, we can compute its value about the desired axis by means of the following relation, first derived by Joseph Louis Lagrange (Fig. 53).

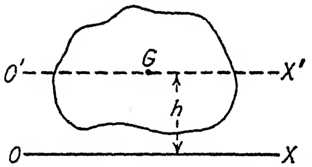


FIG. 53. Moments of Inertia about Parallel Axes

If I_A is the moment of inertia of the body about any axis OX;
I_G is the moment of inertia of the body about a parallel axis $O'X'$ through the center of gravity G of the body (Sec. 73);
M is the total mass of the body; and
h is the distance between parallel axes;

Lagrange's theorem * is:

$$I_A = I_G + Mh^2. \tag{67}$$

Since in this equation Mh^2 is always positive, it follows that for all axes having a given direction the smallest moment of inertia is that about the axis through the center of mass.

VALUES OF MOMENT OF INERTIA FOR REFERENCE

Body	Position of Axis	Moment of Inertia
Uniform thin rod of length l	Normal to length at center	$M\frac{l^2}{12}$
Uniform thin rod of length l	Normal to length at end	$M\frac{l^2}{3}$
Rectangular sheet of sides a and b	Through center normal to sheet	$M\frac{a^2+b^2}{12}$
Thin circular sheet of radius r	Through center normal to sheet	$M\frac{r^2}{2}$
Thin circular sheet of radius r	Along a diameter of circle	$M\frac{r^2}{4}$
Solid sphere of radius r	Any diameter	$\frac{2}{5}Mr^2$

59. Radius of gyration. The radius of gyration K of a body with respect to a given axis is defined as the number which, squared and multiplied by the total mass M of the body, will yield the moment of inertia I of the body about the given axis. Algebraically,

* For derivation, see N. C. Riggs, *Applied Mechanics for Engineers* (New York, The Macmillan Company, 1930), Sec. 64.

$$K \equiv \sqrt{\frac{I}{M}}$$

from which

$$I = K^2M. \tag{68}$$

Thus, if a flywheel has a mass of 500 lb and a radius of gyration of 2 ft about a certain axis of rotation, its moment of inertia about that axis is:

$$I = (2)^2 \times 500 = 2000 \text{ lb-ft}^2.$$

Solved Problem

A string is wound around a circular disk mounted to turn on an axis perpendicular to the disk at its center. The disk weighs 16 lb and is 12 in. in diameter. Neglecting friction and the mass of the string, what force pulling on the string will give the disk a speed of 120 rpm in 6 sec? (See Fig. 54.)

Known:

$$\begin{aligned} M &= 16 \text{ lbm} \\ r &= 6 \text{ in.} = 0.5 \text{ ft} \\ \omega_2 &= 120 \text{ rev/min} = 2 \text{ rev/sec} = 12.57 \text{ rdn/sec} \\ \omega_1 &= 0, \qquad t = 6 \text{ sec.} \end{aligned}$$

Required:

$$F = ?$$

FIRST SOLUTION

From Sec. 58,

$$I = M\frac{r^2}{2} = 16 \text{ lbm}\frac{(0.5 \text{ ft})^2}{2} = 2 \text{ lbm-ft}^2$$

$$\alpha = \frac{\omega_2 - \omega_1}{t} = \frac{12.57 \text{ rdn/sec}}{6 \text{ sec}} = 2.1 \text{ rdn/sec}^2$$

$$Fr = \mathfrak{T} = I\alpha. \qquad \text{by Eq. (65)}$$

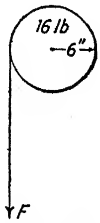


FIG. 54

Substituting these values in Eq. (65),

$$F(0.5 \text{ ft}) = (2 \text{ lbm-ft}^2)\left(2.1\frac{1}{\text{sec}^2}\right) = 4.2\frac{\text{lbm-ft}^2}{\text{sec}^2}\,^*$$

$$\begin{aligned} F &= \frac{4.2\frac{\text{lbm-ft}^2}{\text{sec}^2}}{0.5 \text{ ft}} = 8.4 \text{ lbm}\frac{\text{ft}}{\text{sec}^2} \\ &= 8.4 \text{ poundals} \\ &= \frac{8.4}{32} \text{ lbf} \\ &= 0.26 \text{ lbf.} \end{aligned}$$

* The radian disappears from the expression because, angle being a pure number, its unit is the pure number (1).

SECOND SOLUTION

Exactly the same result is obtained if we use British engineering units. The given quantities are the same as above except that M must be reduced to *slugs*.

$$M = 16 \text{ lbm} = \tfrac{16}{32} \text{ slug}$$

$$I_C = (0.5 \text{ slug}) \frac{(0.5 \text{ ft})^2}{2} = \frac{1}{16} \text{ slug-ft}^2.$$

Substituting these values in Eq. (65):

$$F(0.5 \text{ ft}) = \left(\frac{1}{16} \text{ slug-ft}^2\right)\left(2.1 \frac{1}{\text{sec}^2}\right)^*$$

$$F = \frac{4.2}{16} \frac{\text{slug-ft}^2}{\text{ft-sec}^2} = 0.26 \text{ slug-}\frac{\text{ft}}{\text{sec}^2}$$

$$= 0.26 \text{ lbf}.$$

60. Centripetal force. According to Newton's first law of motion, a particle P (Fig. 55), having a velocity v, will continue in a straight line unless acted upon by some external force.

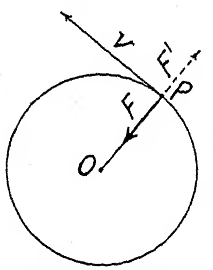


FIG. 55. Centripetal Force

Hence, if we wish P to describe a circle of radius r, we must apply a force F to keep it on the circle. In Sec. 33 it was shown that in order for a particle to describe a circle with uniform speed it must have an acceleration toward the center (centripetal) of

$$a = \frac{v^2}{r}.$$

If the mass of the particle is M, the **centripetal force** necessary to produce this acceleration is, by Newton's second law:

$$F = M\frac{v^2}{r} \tag{69}$$

provided absolute or B.E. units are used.

If P is attached to the axis O by means of a string, F is the force the string must exert upon P.

The particle P will react against the string with a force F' equal and opposite to F in accordance with Newton's third law. This inertia reaction, equal and opposite to centripetal force, is often called **centrifugal force.**

It should be clearly understood, however, that centrifugal force does *not* cause a body to fly away from the center. When the string is broken, the force F no longer acts on the particle; hence the

reaction F' also vanishes, and the particle then continues along the tangent in accordance with Newton's first law, because it is now acted upon by no force.

61. Relativity of motion may be well illustrated here.

Consider a man fastened on a horizontal table so that he is unconscious of the motion of the table, which rotates about a vertical axis with uniform angular velocity. Let a stone near the circumference of the table be restrained from sliding off by a string which the man holds in his hand. The man will feel on his hand a constant pull, which is the inertia reaction of the stone, i.e., the centrifugal force. If he releases the stone, it appears to him to fly off along a radius, because he is turning with the table.

A bystander, however, whose frame of reference is fixed to the earth, perceives the motion of the table and the man, and the stone is seen to fly off along a tangent.

The man on the table would report that the stone moved off under the action of a radial force outward. The bystander would report that the stone moved off under the action of **no force,** for it continued to move in a straight line with the constant velocity (tangential) which it had at the moment it was released—and according to Newton's first law this is what a body does under the action of no force.

Each observer is correct relative to the frame of reference in which he is at rest.

62. The gyroscope. Analogous to the uniform linear motion of a body around a circle, we have **in angular motion** the uniform precessional motion of a gyroscope.

Strictly speaking, every rotating body is a gyroscope. For experimental purposes, however, a gyroscope usually consists of a heavy wheel G (Fig. 56a) mounted in a gimbal ring so as to spin freely about its axis AB. If the wheel is spun at high speed, it exhibits **angular inertia** (i.e., resistance to change of the direction of its axis of spin) in a striking way, so that if supported at one end only of its axis it seems to defy gravity. But as the following analysis shows, its motion is just what would be expected.

Let the gyroscope be supported at one end on a pedestal, as shown, so that it is free to turn about the three axes AB, XY, and JK.

The weight W of the wheel and the reaction of the support produce a torque, $\mathfrak{T} = Wl$, about the axis XY. This torque produces an angular acceleration α parallel to $\mathfrak{T}$.

On account of this acceleration, there will be a change αt in the angular velocity ω of the wheel in the extremely small time t. This change αt is a vector parallel to the vector $\mathfrak{T}$ and therefore normal to ω_1, so that when it is added vectorially (Fig. 56b) to ω_1 we get a new angular velocity ω_2.

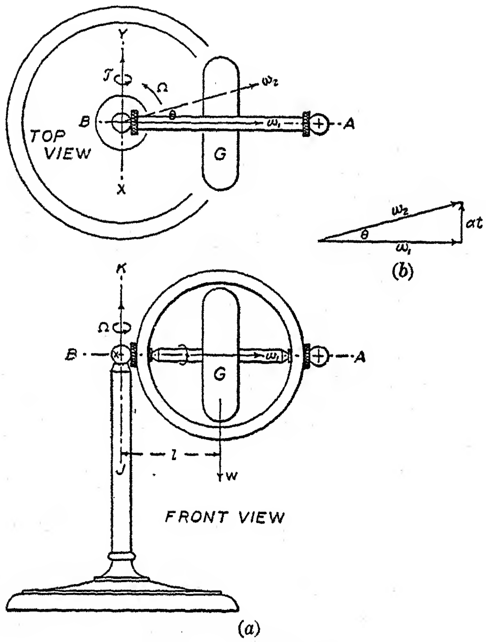


FIG. 56. Gyroscope.

Since t is extremely small, θ is an extremely small angle and its chord αt may be considered equal to its corresponding arc, and $\omega = \omega_1 = \omega_2$, numerically. Hence, by Eq. (2)

$$\theta = \frac{\alpha t}{\omega}$$

whence

$$\frac{\alpha}{\omega} = \frac{\theta}{t} \equiv \Omega \qquad (70)$$

where Ω, which is called the **angular velocity of precession,** is the time rate of change of θ; i.e., the angular velocity with which the spin axis AB turns toward the direction of ω_2 about the axis JK.

Thus theory predicts that the gyroscope, instead of falling off the support as would seem naturally necessary, should precess with an angular velocity Ω about the vertical axis JK, the axis of precession. Experiment shows that this is just what it does.

The torque necessary to produce precession is readily found. By Newton's second law for rotation, $\mathfrak{T} = I\alpha$, where I is the moment of inertia of the moving mass. Hence $\alpha = \mathfrak{T}/I$, and on substituting this in Eq. (70) we obtain

$$\frac{\mathfrak{T}}{I\omega} = \Omega$$

whence

$$\mathfrak{T} = I\omega\Omega. \tag{71}$$

That this precessional torque of the gyroscope is exactly analogous to centripetal force will be seen from the following analysis of the uniform motion of a body B around a circle (Fig. 57).

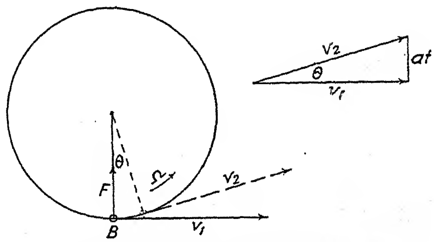


FIG. 57. Uniform Circular Motion

Here v_1 and v_2 correspond to ω_1 and ω_2, respectively, and M of the body B corresponds to I of the gyroscope. Taking the mathematical steps in exactly the same order as before,

$$\theta = \frac{at}{v}$$

$$\frac{a}{v} = \frac{\theta}{t} \equiv \Omega.$$

By Newton's second law for translation, $F = Ma$; whence $a = \dfrac{F}{M}$, so that

$$\frac{F}{Mv} = \Omega$$

and

$$F = Mv\Omega \tag{72}$$

which is exactly analogous to Eq. (71), F corresponding to $\mathfrak{T}$.

By Eq. (34),

$$\Omega = \frac{v}{r}$$

which when substituted in Eq. (72) gives:

$$F = M\frac{v^2}{r}.$$

This is the same as the expression found for centripetal force in Sec. 60.

Hence the uniform precessional motion of the gyroscope is in angular motion analogous to the uniform linear motion of a body around a circle.

Equation (71) shows that if no torque $\mathfrak{T}$ is applied to the gyroscope, there will be no change Ω in the direction of its axis of spin. Hence a gyroscope mounted in three gimbal rings and **kept at high speed** will maintain the direction of its axis quite accurately if the friction at the pivots is slight.

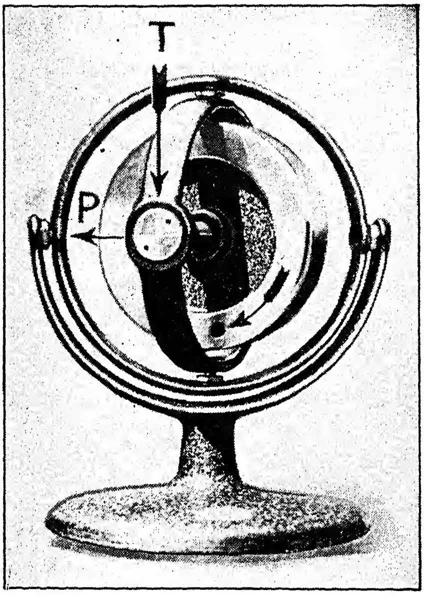


Fig. 58. Sperry Gyroscope. (Courtesy Sperry Gyroscope Co., Inc.)

If such a gyroscope (Fig. 58) is set on a table with its spin axis horizontal and east and west, and its speed is maintained electrically, it will show the rotation of the earth, for after six hours—the earth having turned through 90°—the axis will appear vertical.

On account of this property of directional inertia, the gyroscope is widely used to steer airplanes, ships, and torpedoes. Dr. Elmer E. Sperry utilized this property and the phenomenon of precession to construct the gyroscopic compass (Fig. 59), in which a gyroscope 10 in. in diameter and rotating 6000 rpm is caused to precess until its axis is parallel to the axis of the earth. It then maintains that direction and may be used to steer a ship automatically.

Dr. Sperry has also successfully employed gyroscopes with enormous rotors (100 tons each) to stabilize ships. Such gyrostabilizers reduce the roll of a ship from 26° to 3°.

PROBLEMS

1. What force acting continuously will give a truck which weighs 2880 lb a speed of 15 mph on a level track in 33 sec, friction being neglected?

2. If an auto weighs 2560 lb, what force is necessary to give it a speed of 30 mph in 11 sec on a level road, neglecting friction?

3. If the piston and crosshead of a steam engine together have a mass of 200 lb, what force will be required to give them a speed of 30 ft/sec in 5 sec, considering the guides horizontal and frictionless?

4. A mass of 32 lb is drawn along a frictionless horizontal plane by a mass

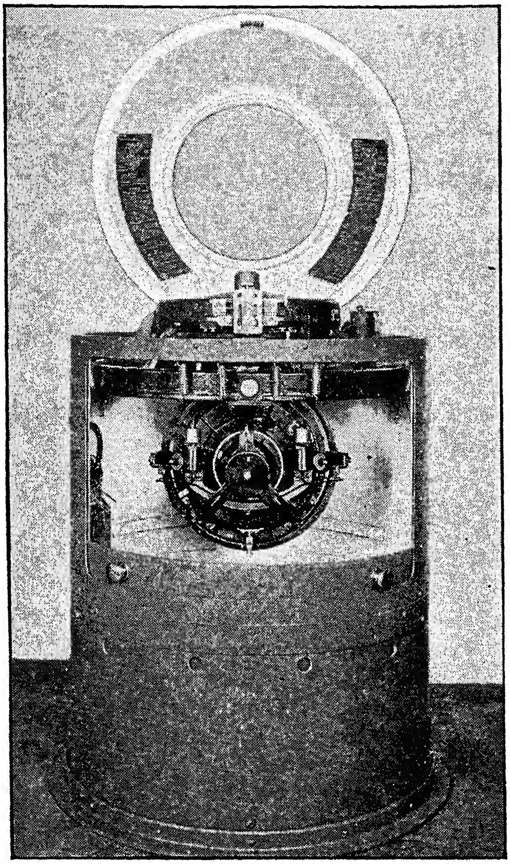


FIG. 59. Sperry Gyroscopic Compass.
(Courtesy Sperry Gyroscope Co., Inc.)

of 8 lb hanging vertically over a massless, frictionless pulley. Find the acceleration of the system and the tension in the string.

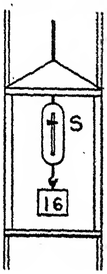
5. A mass of 80 lb is drawn along a frictionless horizontal plane by a mass of 16 lb hanging vertically over a massless, frictionless pulley. Find the acceleration of the system and the tension in the string.

6. Two bodies weighing 60 and 40 gm, respectively, are suspended by a massless cord over a massless, frictionless pulley. Find the acceleration of the system and the tension in the cord.

7. Two masses of 50 and 30 lb, respectively, are hung by a massless cord over a massless, frictionless pulley. Find the acceleration of the system and the tension in the cord.

8. A body whose mass is 16 lb is suspended in an elevator by means of a spring balance. What will the balance read when the elevator is ascending

with a uniform velocity; ascending with an acceleration of 8 ft/sec^2; descending with an acceleration of 8 ft/sec^2?



PROB. 8

9. The car of an elevator weighs 1600 lb and is partially counterbalanced by a weight of 1280 lb. What is the tension in the supporting rope when the car is moving with a constant speed of 10 ft/sec? When its acceleration is 10 ft/sec^2 upward? When its acceleration is 10 ft/sec^2 downward?

10. Two men, each pushing with a force of 150 lb on a wagon, are just able to move it. With what force would each have to push if he applied it at the top of the wheels?

11. A ball 10 cm in diameter, having a mass of 200 gm, is mounted with its center at one end of a rod 1 meter long having a mass of 100 gm. Find the moment of inertia about an axis through the other end of the rod.

12. A flywheel has a mass of 1000 lb and a radius of gyration of 2 ft. Find the angular acceleration produced by a torque of 40 lb-ft.

13. A flywheel weighs 2000 lb and has a radius of gyration of 3 ft. What torque will be required to bring the wheel uniformly to a speed of 240 rpm in 40 sec? How many revolutions will be made in coming up to this speed?

14. The rotor of a steam turbine has 120 buckets whose mean distance from the axis is 10 in. If the tangential force of the steam on each bucket is 0.6 lb and the moment of inertia of the rotor is 25,600 lb-ft^2, what is the angular acceleration and how long will it take to come up to a speed of 1800 rpm?

15. The rotor of a motor has 150 conductors on a cylinder 50 cm in diameter. If each conductor is acted upon by a force of 40,000 dynes and the moment of inertia of the rotor is 1,500,000 gm-cm^2, what is its angular acceleration, and how long will it take to come up to a speed of 3600 rpm?

16. A motor has in its rotor 100 conductors mounted 30 cm from and parallel to the axis. If the force on each conductor is 50,000 dynes, what will be the angular acceleration of the rotor if its moment of inertia is 30,000,000 gm-cm^2 How long will it take to come up to a speed of 2500 rpm?

17. A car weighing 1.5 tons takes a level curve of 100-ft radius at 45 mph. What force tends to overturn it?

18. What force tends to overturn a 3-ton truck which turns a corner on a circle of 50-ft radius at 20 mph?

19. If the ends of a speedway have radii of 200 ft, what force tends to overturn a racing car whose mass is 3000 lb when its speed is 120 mph?

20. A ball weighing 2 lb is whirled in a vertical circle at the end of a string 3 ft long so that it makes 400 rpm. What is the tension on the string at the top and at the bottom of its path, respectively?

21. A mass of 2 lb is whirled in a vertical circle at the end of a string 3 ft long whose tensile strength is 5 lb of force. At what speed in rpm will the string break?

22. A beaker of water is whirled in a vertical circle of 4-ft radius. What is the least speed in rpm at which the water will not spill out?

23. A boy weighing 80 lb sits on a "joy wheel" 3 ft from the axis of rotation.

If the resistance due to friction is 20 lb, at what angular speed of the wheel will he begin to slide?

24. A mass M is suspended by a helical spring whose mass may be neglected, show that if it is set in vibration vertically the period will be $2\pi\sqrt{M/k}$, where k is the force required to elongate the spring unit length. (See Sec. 175.)

25. A vertical helical spring whose mass is neglected is stretched 3 in. by a weight of 2 lb. What will be the period of oscillation of an 8 lb body suspended by the spring?

26. A helical spring whose mass may be neglected is stretched 2 in. when a mass of 4 lb is suspended by it. What is the period of vibration? What is the maximum restoring force when the spring is elongated an additional 3 in. and then released? What is the kinetic energy when the mass passes through the point of equilibrium? What is then the speed of the body?

27. A mass of 50 gm is supported by a helical spring whose mass may be neglected. If a mass of 10 gm elongates the spring an additional 5 cm, what is the period of the spring when loaded with 50 gm?

28. A mass of 10 gm is suspended on a spring so that when displaced from the equilibrium position it vibrates up and down with an amplitude of 4 cm and a frequency of 360 vibrations per minute. Find the tension of the spring at the top of the path, at the center of the path, and 1 cm from the bottom of its path.

CHAPTER V

STATICS

63. Rest and equilibrium. A body is said to be in a state of rest when its velocity is zero.

$$\text{Rest:}\quad \left.\begin{aligned} v &= 0 \quad \text{for translation} \\ \omega &= 0 \quad \text{for rotation.} \end{aligned}\right\} \tag{73}$$

A body is said to be in a state of **equilibrium** when its acceleration is zero.

$$\text{Equilibrium:}\quad \left.\begin{aligned} a &= 0 \quad \text{for translation} \\ \alpha &= 0 \quad \text{for rotation.} \end{aligned}\right\} \tag{74}$$

A body may be at rest without being in equilibrium; or it may be in equilibrium without being at rest. The long vibrating pendulum of Fig. 60 illustrates this well.

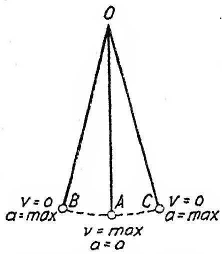


FIG. 60. Long Pendulum

When the bob passes through the point A, its velocity is a maximum and its acceleration is zero. At the end positions B and C, however, its velocity is zero and its acceleration maximum.

Hence at A the bob is in equilibrium but not at rest; while at B and C it is at rest but not in equilibrium.

These facts may easily be verified by means of the relations in Eq. (42), since the motion of a pendulum closely approximates simple harmonic motion, if it is remembered that for these relations $t = 0$ at C and that x is measured from A.

The **equilibrant** of a given system of coplanar forces is a single force or a couple that will hold in equilibrium a body which is acted upon by the given system of forces. For forces not all in the same plane, the equilibrant must be a single force and a couple whose axis is parallel to the single force.

The **resultant** of a given system of coplanar forces is a single

force or a couple that will produce the **same result** as is produced by the given system.

Since the equilibrant would equilibrate the original system, it must then equilibrate the resultant. But a single force can be equilibrated, or balanced out, only by an equal and opposite force having the same line of action; and a couple, by an equal and opposite couple.

Hence, **the equilibrant of a system of forces is always equal and opposite to the resultant;** and if a single force, it has the same line of action as the resultant.

64. Statics is that part of dynamics which deals with forces in equilibrium. In many instances the bodies upon which the forces act are both in equilibrium and at rest, but it is not necessary that they should be at rest. The results are equally true for bodies in uniform motion.

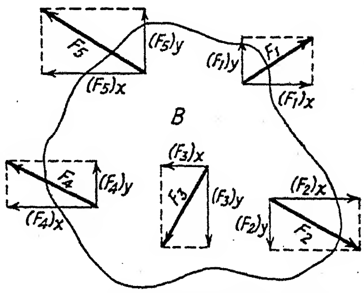


FIG. 61. Equilibrium—General Conditions

General conditions of equilibrium. Consider a body B (Fig. 61), held in equilibrium by *any* system of forces $F_1, F_2, \ldots F_n$. Let each force be replaced by its horizontal and vertical components (Sec. 25) as shown.

If the body is in equilibrium with regard to translation in a vertical direction, **the sum of the forces upward must equal the sum of the forces downward.** For otherwise, one or the other set of forces must be the greater, say, those upward. These would then give the body an acceleration upward, which is contrary to the hypothesis that the body is in equilibrium.

Writing this fact as an equation,

$$\text{Forces upward} = \text{Forces downward}$$

$$(F_1)_y + (F_4)_y + (F_5)_y = (F_2)_y + (F_3)_y. \tag{a}$$

Transposing,

$$(F_1)_y - (F_2)_y - (F_3)_y + (F_4)_y + (F_5)_y = 0. \tag{b}$$

Examination of Eq. (b) shows the left side to be the algebraic sum of the vertical components of the forces, i.e., the sum of the

vertical components giving to each its proper algebraic sign. Hence the equation may be written in the briefer form:

$$\Sigma F_y = 0. \tag{c}$$

Similarly, if the body is in equilibrium with regard to translation horizontally,

$$\text{Forces to right} = \text{Forces to left}$$

$$(F_1)_x + (F_2)_x = (F_3)_x + (F_4)_x + (F_5)_x.$$

Transposing,

$$(F_1)_x + (F_2)_x - (F_3)_x - (F_4)_x - (F_5)_x = 0 \tag{d}$$

or,

$$\Sigma F_x = 0. \tag{e}$$

In the case of motion in space, we would take a third, or Z-, axis perpendicular to the X- and Y-axes; it would then be necessary also that

$$\Sigma F_z = 0. \tag{f}$$

If Eqs. (c), (e), and (f) are satisfied, the body will be **in equilibrium as regards translation.** It might still have acceleration in rotation, i.e., in spinning about an axis.

For rotation, if there is to be equilibrium, there must obviously be no outstanding torque; for if there were it would produce an angular acceleration in accordance with Sec. 56, which is contrary to the hypothesis that the body is in equilibrium. Hence,

$$\text{Clockwise torques must} = \text{Counterclockwise torques.}$$

That is

$$\Sigma \mathfrak{T} = 0 \qquad \left(\begin{matrix}\text{about any axis}\\ \text{one may choose}\end{matrix}\right).$$

This will be true if $\Sigma \mathfrak{T} = 0$ about each of three concurrent axes not in the same plane.

Collecting these conditions, we have the **general conditions of equilibrium:**

$$\left.\begin{aligned}\Sigma F_x &= 0\\ \Sigma F_y &= 0\\ \Sigma F_z &= 0\end{aligned}\right\} \quad \text{for equilibrium in translation.} \tag{75}$$

$$\left.\begin{aligned}\Sigma \mathfrak{T}_x &= 0\\ \Sigma \mathfrak{T}_y &= 0\\ \Sigma \mathfrak{T}_z &= 0\end{aligned}\right\} \quad \text{for equilibrium in rotation.} \tag{76}$$

65. Equilibrium of nonconcurrent forces. The conditions of Sec. 64 must be satisfied whenever a system of forces is in equilibrium. The way in which these conditions may be employed is illustrated in the solution of the following problem, in which, for simplicity, the forces are taken all in the same plane.

Solved Problem

Given a body B (Fig. 62) acted upon by the forces:

$F_1 = 3$ gmf $\qquad F_3 = 4$ gmf
$F_2 = 6$ gmf $\qquad F_4 = 5$ gmf

whose points of application and directions are shown by the squares, 1 cm on a side, which have been laid off on the body; it is required to find the equilibrant of the given system of forces.

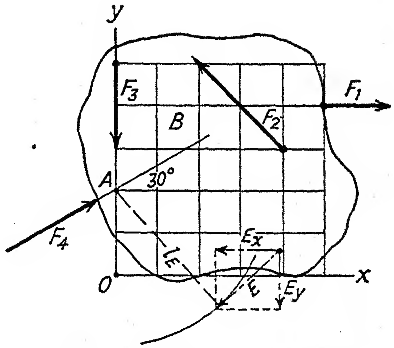


Fig. 62. Noncurrent Forces—First Solution

Analytical Solution

Let E be the value of the equilibrant which is to be determined, and consider it to have been found and to be acting on the body. The latter will then be in equilibrium, and hence the algebraic sum of the torques T about any axis will be zero. This being true for any axis, we will take the axis at a convenient place, say, perpendicular to the paper at point A. Each force is then resolved into its vertical F_y and horizontal F_x components as in Sec. 25; and its lever arm l and torque T with respect to A are computed. Counterclockwise torques are called + and clockwise −, as is customary.

The solution is condensed by arranging these values in a table with the angles θ which the forces make with the X-axis, and the sines and cosines of these angles, as shown.

Force		Angle θ	Cos θ	Sin θ	Lever Arm l	F_x	F_y	Torque $\mathfrak{T} = Fl$
F_1	3	0	1	0	2	3	0	−6.0
F_2	6	135	−0.707	0.707	3.54	−4.24	4.24	+21.22
F_3	4	−90	0	−1	0	0	−4.0	0
F_4	5	30	0.866	0.5	0	4.33	2.5	0
E	?	?	?	?	?	E_x	E_y	$\mathfrak{T}_E$

From the last three columns of the table we can write down the conditions of equilibrium at once:

$$\Sigma F_x = 3 - 4.24 + 0 + 4.33 + E_x = 0; \qquad E_x = -3.09 \text{ gmf}$$
$$\Sigma F_y = 0 + 4.24 - 4.0 + 2.5 + E_y = 0; \qquad E_y = -2.74 \text{ gmf}$$
$$\Sigma \mathfrak{T} = -6 + 21.22 + 0 + 0 + \mathfrak{T}_E = 0; \qquad \mathfrak{T}_E = -15.22 \text{ cm-gmf}$$
$$E = \sqrt{E_x^2 + E_y^2} = \sqrt{(-3.09)^2 + (-2.74)^2} = 4.13 \text{ gmf}$$
$$\tan \theta_E = \frac{E_y}{E_x} = 0.887; \quad \theta_E = 221°34'.$$

Here we have the magnitude and direction of the equilibrant E.

To determine the line of action, we note that the torque of E is:

$$\mathfrak{T}_E = -15.22 \text{ cm-gmf.}$$

Hence the lever arm of E must be:

$$l_E = \frac{-15.22 \text{ cm-gmf}}{4.13 \text{ gmf}} = -3.69 \text{ cm}$$

which, being negative, means that it must be so taken that the torque will be clockwise.

Thus the line of action of the equilibrant is tangent to a circle of radius 3.69 cm having its center at A. The equilibrant must act downward and to the left, but may be considered to act anywhere along this tangent line.

From the above solution, it will be seen that each condition of equilibrium gives an equation; and in general there will be enough of these equations to determine the equilibrant in magnitude, direction, and line of action. In this case, since the forces lay all in the same plane, it was not necessary to use the condition, $\Sigma F_z = 0$. Or, more accurately, $\Sigma F_z = 0$ was satisfied because all the components parallel to the Z-axis were zero by hypothesis.

Graphical Solution

Another method of finding the equilibrant of a system of **nonconcurrent forces** is first to find the resultant of the system by successive applications of the law of the parallelogram of vectors (Sec. 21). The equilibrant is then equal and opposite to that resultant with the same line of action.

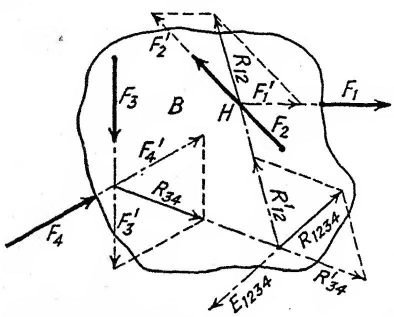


Fig. 63. Noncurrent Forces—Second Solution

The foregoing problem is solved in this manner in Fig. 63. The lines of action of any two of the forces, say, F_1 and F_2, are produced until they meet at some point H. Their values are laid off from this point of intersection as F_1' and F_2', and their resultant R_{12} is found by completing the parallelogram. Similarly, R_{34}, the resultant of F_3 and F_4, is found. Since R_{12} and R_{34} are partial resultants, together they will produce the same result as would the original system. Hence we may consider them as replacing the given forces. The lines of action of these

partial resultants are prolonged until they meet; R_{12} and R_{34} are moved along their lines of action to the positions R_{12}' and R_{34}', respectively; and another parallelogram of forces is constructed giving the complete resultant R_{1234} in magnitude, direction, and line of action. The equilibrant required is then the force E_{1234}, equal and opposite to this resultant. It will be seen to be the same as the result of the first solution.

This graphical method may obviously be applied to any number of forces, slipping them along their lines of action to the point of intersection, and combining them and their partial resultants two at a time by vector addition, until only a single force or a couple remains. This is the resultant of the system in magnitude, direction, and line of action. By taking a force or couple equal and opposite to this, we have the equilibrant. If the drawing is carefully done, the method is sufficiently accurate for most engineering purposes.

It may be asked why we do not use the polygon of forces to solve the above problem. It will be recalled that **the polygon of forces does not give the line of action.**

Three special cases of force systems at once suggest themselves.

66. Case I: When the forces are concurrent, i.e., all meet in a point.

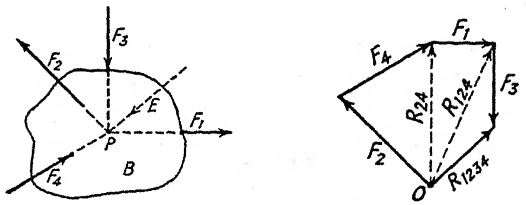
Given the system of forces shown in Fig. 64; required the equilibrant of the system.

The *analytical solution* when the forces are concurrent is precisely the same as in the preceding section, where they are nonconcurrent, except that since the line of action of the equilibrant must obviously pass through the common point of the given lines of action, it is not necessary to use the condition, $\Sigma \mathfrak{Z} = 0$. For the same reason the graphical solution also is simplified.

Graphical solution. As before, we first determine the resultant of the forces and then reverse it. For the sake of clearness, it is usual to draw the vectors representing the given forces in their proper places on the body B upon which they act (**space diagram**); and then to redraw them in a separate figure (**force diagram**) when constructing the force polygon.

Accordingly, let Fig. 64a represent any system of forces acting upon B such that their lines of action, if produced, will meet in a point P.

The forces being vectors, we then determine their resultant as in Sec. 22. Beginning at some point O, Fig. 64b, draw F_2, F_4, F_1, F_3, equal and parallel to the corresponding forces in the space diagram, using any convenient scale. The order in which the forces are taken is quite immaterial, but it is **essential** that



SPACE DIAGRAM (a)

FORCE DIAGRAM (b)

Fig. 64. Concurrent Forces

the beginning of each vector shall be the end (arrow point) of the preceding vector.

The forces then form a polygon whose diagonals are the partial resultants, and whose closing side is the **complete resultant** (R_{1234}) as in Sec. 22, its arrow pointing away from the origin.

The **equilibrant** also will be represented by the closing side of the polygon, but with its arrow pointing the same way around the figure as do the arrows of the components, since it is equal and opposite to the resultant. This vector is then transferred (with dotted line) to the space diagram, where it acts through the common point P.

The force polygon is not restricted to a single plane, but may be a figure twisted in space when the problem is in three dimensions.

Corollary. The force polygon gives the resultant of a system of nonconcurrent forces also, in magnitude and direction, **but does not give its line of action.** The forces of Fig. 64a were taken the same in magnitude and direction as those in Fig. 63 in order to illustrate this fact. The resultants in the two cases will be seen to be exactly equal in magnitude and parallel in direction—but we would not know where to place the line of action of E in Fig. 63 after finding its magnitude and direction in the force diagram of Fig. 64b, for these forces have no common point. There is a

method for determining this line of action by constructing in the space diagram an "equilibrium, or funicular polygon."*

67. Case II: When a body is in equilibrium under the action of three forces.

Given the body B, Fig. 65a, in equilibrium under the action of the forces, F_1, F_2, and F_3; required the relations among the forces.

The general conditions of equilibrium (Sec. 64), apply to this case as to all others, but certain consequences of these conditions are easily borne in mind and simplify the determination of the third force when the other two are known. As this is one of the

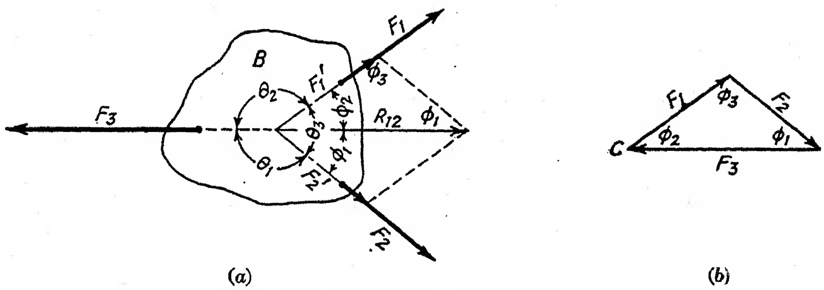


Fig. 65. Equilibrium—Three Forces

most common problems met in practice, the facts should be noted carefully.

Let the resultant R_{12} of F_1 and F_2 be found as in the graphical solution of Sec. 65. Since a resultant is equivalent to its component forces, we may consider F_1 and F_2 to be removed, and let R_{12} act in their stead. Then, since F_3 held F_1 and F_2 in equilibrium, it will now hold R_{12} in equilibrium. To do this it must be equal in magnitude, opposite in direction, and must have the same line of action as R_{12}. Therefore, it must lie in the plane of F_1 and F_2 and must pass through their intersection; i.e., the three forces must meet in a point.

If we draw the force polygon (Fig. 65b), it should be a closed figure. For F_3 is the equilibrant of F_1 and F_2 and is therefore, by Sec. 66, the closing side of the force polygon. Or, supposing the polygon did not close, then the side necessary to make it close would represent the resultant, with its arrow pointing away from

* Merriman and Jacoby, *A Text-book of Roofs and Bridges* (New York, John Wiley and Sons, 1926), p. 15.

the origin O. But this is contrary to the hypothesis that the system is in equilibrium and has zero resultant. Hence the polygon must close.

Finally, from the law of sines,

$$F_1 : F_2 : F_3 = \sin \varphi_1 : \sin \varphi_2 : \sin \varphi_3$$
$$= \sin (180 - \theta_1) : \sin (180 - \theta_2) : \sin (180 - \theta_3)$$
$$= \sin \theta_1 : \sin \theta_2 : \sin \theta_3. \tag{77}$$

Eq. 77 is the mathematical statement of **Lami's theorem: When a body is held in equilibrium by three forces, each force is proportional to the sine of the angle between the other two.**

Summarizing these facts: If a body is in equilibrium under the action of three forces,

(a) The forces must lie in a plane.

(b) The forces must meet in a point.

(c) The force polygon must be a closed triangle; i.e., the forces must satisfy Lami's theorem.

Solved Problems

1. A stiff rod BD (Fig. 66a), 5 ft long, is attached to a wall at B by a pin-joint. It is supported in a horizontal position by a rope fastened to the rod at C, 3 ft from B, its other end being fastened to the wall at A, 4 ft vertically above B. Neglecting the weights of the rod and the rope, what is the reaction of the wall at B and what is the tension in the rope when a load of 60 lb hangs at D.

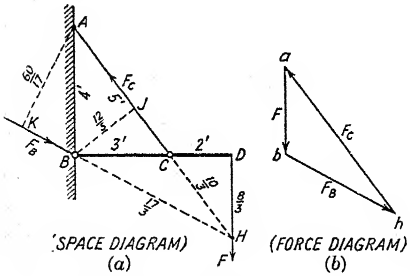


FIG. 66. Shelf Problem

Analysis: Since the weights of the rod and the rope are neglected, the rod BD is held in equilibrium by three forces: the load F at D; the reaction of the wall F_B at B; and the tension in the rope F_C which acts at C.

Hence, from the summary above, these forces must (a) lie in a plane; (b) meet in a point; (c) form a closed triangle.

FIRST SOLUTION (GRAPHICAL)

The force F_C must act along AC, since a flexible rope can transmit only tension in the direction of its length. F_C and F intersect at H. Hence F_B, produced, must also pass through H.

Construct the force diagram by drawing the vector ab parallel to F—and hence parallel also to AB—to represent ($F = 60$ lb) to some convenient scale. Then draw ah parallel to AH and bh parallel to BH. These meet at h and cut off the correct lengths to represent F_C and F_B, respectively.

The values of F_C and F_B in pounds are found by measuring ah and bh in inches and multiplying by 64, since each inch represents 64 lb of force. This gives:

$$F_C = \tfrac{125}{64} \text{ in.} \times 64 \text{ lbf/in.} = 125 \text{ lbf}$$
$$F_B = \tfrac{85}{64} \text{ in.} \times 64 \text{ lbf/in.} = 85 \text{ lbf.}$$

Second Solution (Geometrical)

Since the force triangle abh is similar to the space triangle ABH, their sides being respectively parallel by construction, we have from geometry that

$$\frac{bh}{ab} = \frac{BH}{AB}. \tag{a}$$

To find BH we first obtain DH from the similar triangles CDH and CBA. Here,

$$\frac{DH}{CD} = \frac{AB}{CB} = \frac{4}{3}$$

whence

$$DH = \tfrac{4}{3} \times 2 = \tfrac{8}{3} \text{ ft.}$$

Therefore

$$BH = \sqrt{(5)^2 + (\tfrac{8}{3})^2} = \tfrac{17}{3} \text{ ft.}$$

Putting this value of BH in Eq. (a) with the values of ab and AB, and writing F_B for bh, we have:

$$\frac{F_B}{60} = \frac{\frac{17}{3}}{4}$$

whence

$$F_B = 85 \text{ lbf.}$$

Similarly,

$$\frac{ah}{ab} = \frac{AH}{AB}. \tag{b}$$

To find AH, we first obtain CH.

$$CH = \sqrt{(2)^2 + (\tfrac{8}{3})^2} = \tfrac{10}{3} \text{ ft}$$

so that

$$AH = 5 + \tfrac{10}{3} = \tfrac{25}{3} \text{ ft.}$$

Putting this value of AH in Eq. (b) with the values of ab and AB, and writing F_C for ah, we have:

$$\frac{F_C}{60} = \frac{\frac{25}{3}}{4}$$

whence

$$F_C = 125 \text{ lbf.}$$

Third Solution (by Moments)

The lever arm of F_C about B is BJ. From similar triangles BJC and ABC,

$$\frac{BJ}{BC} = \frac{AB}{AC} = \frac{4}{5}$$

from which

$$BJ = \tfrac{4}{5} \times 3 = \tfrac{12}{5} \text{ ft.}$$

Hence, applying $\Sigma\mathfrak{Z} = 0$, taking moments about an axis perpendicular to the paper at B,

$$F_C \times \tfrac{12}{5} - F \times 5 = 0$$
$$F_C = 60 \times 5 \times \tfrac{5}{12} = 125 \text{ lbf.}$$

Similarly; the lever arm of F_B about A is AK· and from similar triangles AKB and BDH,

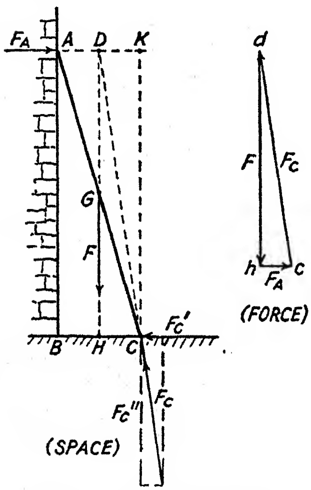


Fig. 67. Ladder Problem

$$\frac{AK}{AB} = \frac{BD}{BH}$$

from which

$$AK = \frac{5}{\frac{17}{3}} \times 4 = \frac{60}{17} \text{ ft.}$$

Hence, applying $\Sigma\mathfrak{Z} = 0$, taking moments about an axis perpendicular to the paper at A,

$$F_B \times \tfrac{60}{17} - F \times 5 = 0$$
$$F_B = 60 \times 5 \times \tfrac{17}{60} = 85 \text{ lbf.}$$

2. A ladder AC (Fig. 67), 25 ft long and weighing 48 lb, rests on the ground at point C, 7 ft from the foot of vertical wall AB. Neglecting friction at A, find the reactions at A and C.

Analysis: Friction at A being neglected, the reaction F_A will be normal to the wall, for a frictionless surface can exert only a normal force. The weight F of the ladder may be considered as acting vertically downward through its center G. The lines of action of these forces intersect at D. The only other force acting on the ladder is the reaction F_C at C.

Since the ladder is in equilibrium under three forces we know that these forces (a) lie in a plane; (b) meet in a point, hence F_C must pass through D, the intersection of F_A and the weight F; and (c) form a closed triangle.

First Solution (Graphical)

Draw the vector dh representing 48 lb to some convenient scale. (In the figure 1 in. = 32 lb.) Draw hc parallel to HC and dc parallel to DC. These lines intersect at c and cut off correct lengths to represent F_A and F_C, respectively. Their values in pounds are found by measuring hc and dc accurately in inches and multiplying by 32, since each inch represents 32 lb. This gives:

$$F_A = 0.22 \text{ in.} \times 32 = 7.04 \text{ lbf}$$
$$F_C = 1.52 \text{ in.} \times 32 = 48.6 \text{ lbf.}$$

The graphical method is usually the simplest, but its accuracy depends entirely upon one's skill in drawing. It is excellent as a rough check upon other methods.

Second Solution (Geometrical)

Known: $F = 48$ lbf
$AB = 24 \text{ ft} = DH$
$BC = 7$ ft.

Computed: $BH = 3.5$ ft
$HC = 3.5$ ft
$DC = 24.25$ ft.

Since the force triangle *dhc* is similar to the space triangle *DHC*, their sides being respectively parallel by construction, we have from geometry;

$$\frac{hc}{dh} = \frac{HC}{DH} \quad \text{or} \quad \frac{F_A}{48} = \frac{3.5}{24}$$

from which

$$F_A = 7 \text{ lbf.}$$

Similarly,

$$\frac{dc}{dh} = \frac{DC}{DH} \quad \text{or} \quad \frac{F_C}{48} = \frac{24.25}{24}$$

from which

$$F_C = 48.5 \text{ lbf.}$$

Third Solution (by Moments)

Another easy solution is found by applying the principle of moments ($\Sigma \mathfrak{T} = 0$), choosing axes so as to give directly either the required forces or their rectangular components.

Thus, taking torques about an axis perpendicular to the paper at C, we have:

$$48 \times 3.5 - F_A \times 24 + F_C \times 0 = 0$$
$$F_A = 7 \text{ lbf.}$$

It is not quite so easy to find F_C because its lever arm about A, for example, would have to be computed. But we can easily find the horizontal component F_C' and the vertical component F_C'' of F_C, which can then be found by adding these components vectorially.

To do this, take torques about an axis perpendicular to the paper at K; then,

$$48 \times 3.5 + F_A \times 0 - F_C' \times 24 + F_C'' \times 0 = 0$$
$$F_C' = 7 \text{ lbf.}$$

And taking torques about an axis perpendicular to the paper at H,

$$48 \times 0 - F_A \times 24 + F_C' \times 0 + F_C'' \times 3.5 = 0$$

$$F_C'' = \frac{24F_A}{3.5} = \frac{24 \times 7}{3.5} = 48 \text{ lbf}$$

from which $F_C = \sqrt{(7)^2 + (48)^2} = 48.5$ lbf.

This last solution brings out clearly the fact that the ladder may be considered in equilibrium under the action of two equal and opposite couples: one, consisting of F and F_C'', being counterclockwise; the other, consisting of F_A and F_C', being clockwise.

Attention is called to the fact that a solution by means of a force polygon alone or by moments alone is possible **only when the forces are concurrent;** in all other cases both must be used.

68. Case III: When a body is held in equilibrium by a system of parallel forces.

Given the body B (Fig. 68) in equilibrium under the action of a system of forces all of which are parallel, as shown; required the conditions of equilibrium for such a system.

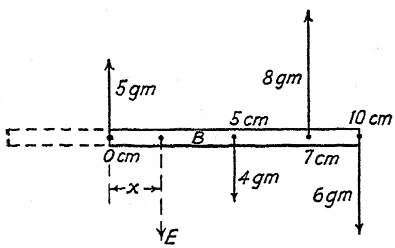


FIG. 68. Parallel Forces

Since the forces are parallel we may take one of our axes (say, the Y-axis) parallel to the forces. Then there will be no components parallel to the X- and Z-axes. Hence,

$$\Sigma F_x = \Sigma F_z = 0$$

and the conditions of equilibrium for parallel forces become

$$\Sigma F_y = 0$$

and

$$\Sigma \mathfrak{I} = 0.$$

Or, since there is only one direction of forces, the subscript y is no longer necessary and the **conditions of equilibrium of parallel forces become:**

$$\Sigma F = 0 \text{ with regard to translation.} \quad (78)$$

$$\Sigma \mathfrak{I} = 0 \text{ with regard to rotation.} \quad (79)$$

Each of these conditions gives us an equation, so that in general we can find two unknowns in a system of parallel forces in equilibrium. These unknowns may be:

1. Magnitude of one force and its line of action (lever arm).
2. Magnitude of one force and line of action of another force.
3. Magnitudes of two forces.
4. Lines of action of two forces.

Solved Problems

1. Find the equilibrant of the system of forces shown in Fig. 68. Assume the equilibrant E, shown dotted, to act at some distance x from the left end of the body. It is required to find the magnitude of E and its line of action as given by x.

Applying the condition, $\Sigma F = 0$, and taking forces upward as + and forces downward as −,

$$5 - E - 4 + 8 - 6 = 0$$
$$E = 3 \text{ gmf, downward.}$$

Applying the condition, $\Sigma\mathfrak{T} = 0$, taking counterclockwise torques + and clockwise torques −, as is customary, and axis at O,

$$5 \times 0 - Ex - 4 \times 5 + 8 \times 7 - 6 \times 10 = 0$$
$$-Ex = 24$$
$$3x = -24$$
$$x = -8 \text{ cm}$$

where the − sign means that x must be laid off to the left of the origin.

Hence the equilibrant is a force of 3 gm acting downward at a distance of 8 cm to the left of the origin. Since the body is not shown extending to the left, the equilibrant must be considered to act on an imaginary extension (dotted).

Had the problem asked for the **resultant** instead of the equilibrant, we should still have found the equilibrant first as shown above. The resultant would then be reported as a force of 3 gm acting **upward** at a distance of 8 cm to the left of O.

The conventions of signs for forces and torques are entirely arbitrary, but those used above are generally adopted. When there is a large number of forces, it may not be easy to determine whether to assume E upward or downward. In this case, if the value of E comes out −, it indicates that the wrong direction was chosen. The magnitude will still be correct. The same is true of the lever arm x, as shown above.

2. A meter-stick weighs 120 gm. A weight of 30 gm hangs at the zero mark, and forces of 60 gm and 90 gm act vertically upward at the 25-cm and 100-cm marks, respectively. Find the resultant of the system. (See Fig. 69.)

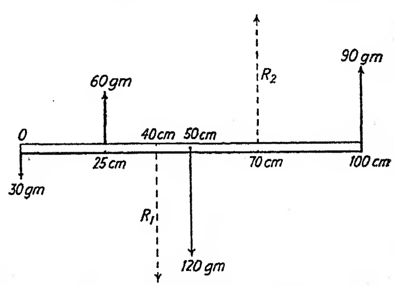


FIG. 69. Resultant Couple

Assume the equilibrant to act upward at a distance x from the 0 on the stick. Applying $\Sigma F = 0$,

$$-30 + 60 - 120 + E + 90 = 0$$
$$E = 0.$$

Since the equilibrant is zero, either the system is already in equilibrium or its equilibrant is a couple. To find which is correct, apply $\Sigma\mathfrak{T} = 0$, taking torques about an axis at 0.

$$30 \times 0 + 60 \times 25 - 120 \times 50 + Ex + 90 \times 100 = 0$$
$$Ex = -4500 \text{ cm-gmf.}$$

Since $\Sigma\mathfrak{T}$ is not zero, the system is *not* in equilibrium and the torque necessary to produce equilibrium is 4500 cm-gm clockwise.

The system would be equilibrated by applying to it *any couple* having a torque of -4500 cm-gm; i.e., there is an infinite number of solutions to the problem.

But there is one particular couple that may have some claim to being **the resultant couple.** It is found by combining all the upward forces into a single force and all the downward forces into a single force, as follows:

(a) Taking the downward forces as a separate system, find their equilibrant E_1 and resultant R_1. Applying $\Sigma F = 0$,

$$E_1 - 30 - 120 = 0$$
$$E_1 = 150 \text{ gmf, upward.}$$

Applying $\Sigma\mathfrak{T} = 0$, taking moments about an axis at 0,

$$30 \times 0 + E_1x_1 - 120 \times 50 = 0$$
$$E_1x_1 = 6000 \text{ cm-gm}$$
$$150x_1 = 6000$$
$$x_1 = 40 \text{ cm.}$$

(b) Taking the upward forces as a separate system, find their equilibrant E_2 and resultant R_2. Applying $\Sigma F = 0$,

$$60 - E_2 + 90 = 0$$
$$E_2 = 150 \text{ gmf, downward.}$$

Applying $\Sigma\mathfrak{T} = 0$, taking moments about 0,

$$60 \times 25 - E_2x_2 + 90 \times 100 = 0$$
$$-E_2x_2 = -10{,}500 \text{ cm-gm}$$
$$150x_2 = 10{,}500$$
$$x_2 = 70 \text{ cm.}$$

This unique couple consisting of E_1 and E_2, acting at 40 cm and 70 cm, respectively, has the torque $\mathfrak{T} = -150 \text{ gm} \times 30 \text{ cm} = -4500$ cm-gm, and is therefore one of the possible equilibrating couples.

$$R_1 = -E_1 = 150 \text{ gmf, downward}$$
$$R_2 = -E_2 = 150 \text{ gmf, upward.}$$

The corresponding resultant couple therefore consists of R_1 and R_2 and has a torque of $+4500$ cm-gm.

69. Center of gravity. In many problems it is necessary to take into account the weight of the body itself. We then need to know the position of the so-called center of gravity of the body.

The **center of gravity** of a body, or system of bodies, is defined as that point at which, if its entire mass were concentrated, the gravitational force on that mass would be the same as before the

concentration. In less general terms, it is the point at which the **weight of a body may be considered concentrated.**

Calculation shows that the center of gravity of a system A with reference to a system B is a different point for different relative positions, or orientations, of A with respect to B. Thus, the center of gravity of the system consisting of the earth and the moon is not the same with reference to the sun when the moon is new and when the moon is in the first quarter phase. A solid or hollow sphere is the only body whose center of gravity remains always the same point.*

Most of our problems deal with the attraction of the earth for a body, i.e., its weight. For bodies of ordinary size, say, not larger than a battleship, the center of gravity is practically a definite point conforming to the above definition, and may be found by several methods.

70. Experimental method of determining the center of gravity of a body. The weight of a body is obviously the resultant of the individual weights of all of its particles. If a body is suspended by a cord at S, it will come to equilibrium when the tension E in the cord is equal and opposite to the weight W of the body (Fig. 70a). Since the direction of weight, or gravity, is that of a plumb line (see Sec. 42); if a plumb bob is hung from the same point S, and the position of the plumb line marked on the body, this mark JK will be the line of action of the weight W.

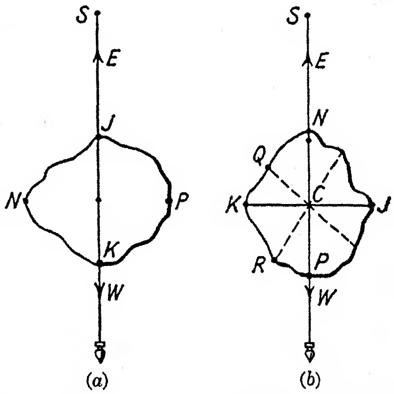


FIG. 70. Center of Gravity—Experimental Determination

Let the body then be suspended from any other point N (Fig. 70b), and the same procedure followed. A second line of action NP of the weight will be secured, which will intersect JK at C.

Suspending the body from as many different points (Q, R, etc.)

* It is said to have taken Newton twenty years to prove this property for the sphere.

as one will, the line of action of the weight will be found in every case to pass through C, which is the center of gravity by the above definition.

The center of gravity is not always a point within or on the body. For example, if an ordinary laboratory tripod is suspended from different points as described above, its center of gravity will be found to be a point in the space between the legs.

Since the weight of a body is always a downward force through the center of gravity, it follows that a body supported at its center of gravity by a vertical force will be in equilibrium in any position. For in every possible position the weight downward and the supporting force upward have the same line of action through the center of gravity, and thus completely neutralize each other.

Another practical method. The above fact suggests that the center of gravity may be found also by the process of balancing a body on a knife edge (sharp edge of a prism) and finding the intersection of two lines on the body along which it will balance.

71. Analytical method of determining the center of gravity of a body. Since the weights of the individual particles of a body are *practically parallel* forces, we can calculate the position of the center of gravity by finding the line of action of the resultant of these little forces by the method of Sec. 68.

In Fig. 71a, call the masses of the particles $m_1, m_2, \ldots m_n$; their weights, $w_1, w_2, \ldots w_n$, respectively; and their equilibrant, E. Imposing the conditions for equilibrium of parallel forces, we have:

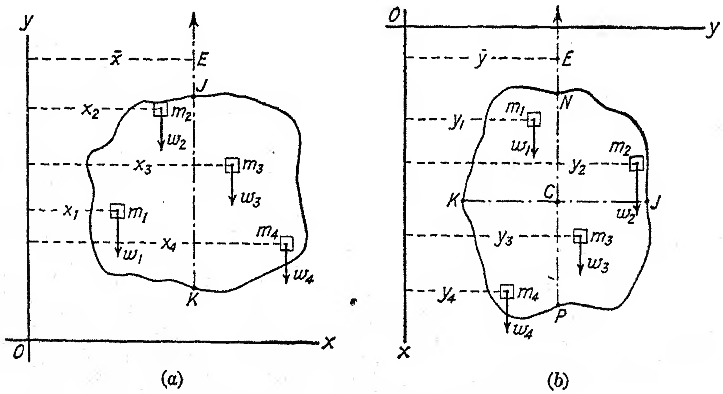


FIG. 71. Center of Gravity

(1) $\Sigma F = 0$:

$$E - w_1 - w_2 \cdot \cdot \cdot - w_n = 0$$
$$E = w_1 + w_2 \cdot \cdot \cdot + w_n \equiv \Sigma w. \qquad \text{(a)}$$

Referring the particles to a convenient set of rectangular axes, and taking the axis of torques at the origin, apply the second condition of equilibrium.

(2) $\Sigma \mathfrak{I} = 0$: axis at origin,

$$E\bar{x} - w_1x_1 - w_2x_2 \cdot \cdot \cdot - w_nx_n = 0$$
$$E\bar{x} = w_1x_1 + w_2x_2 \cdot \cdot \cdot + w_nx_n \equiv \Sigma wx \qquad \text{(b)}$$

from which, by Eq. (a),

$$\bar{x} = \frac{\Sigma wx}{E} = \frac{\Sigma wx}{\Sigma w}. \qquad \text{(c)}$$

If we turn Fig. 71a, axes and all, through 90° clockwise, we get Fig. 71b, in which the weights of the particles, being always vertically downward, are now parallel to the X-axis.

Again applying the conditions of equilibrium,

(1) $\Sigma F = 0$, whence

$$E = \Sigma w \qquad \text{(d) exactly as before.}$$

(2) $\Sigma \mathfrak{I} = 0$: axis again at the origin,

$$E\bar{y} - w_1y_1 - w_2y_2 - \cdot \cdot \cdot - w_ny_n = 0$$
$$E\bar{y} = w_1y_1 + w_2y_2 + \cdot \cdot \cdot + w_ny_n \equiv \Sigma wy$$
$$\bar{y} = \frac{\Sigma wy}{E} = \frac{\Sigma wy}{\Sigma w}. \qquad \text{(e)}$$

Similarly, for a three-dimensional body:

$$\bar{z} = \frac{\Sigma wz}{E} = \frac{\Sigma wz}{\Sigma w}. \qquad \text{(f)}$$

Since these coordinates $(\bar{x}, \bar{y}, \bar{z})$ define the position of the center of gravity for any body, they are taken as the analytical, or quantitative, definition of the center of gravity.

Defining equations for center of gravity:

$$\left.\begin{aligned} \bar{x} &\equiv \frac{\Sigma wx}{\Sigma w} \\ \bar{y} &\equiv \frac{\Sigma wy}{\Sigma w} \\ \bar{z} &\equiv \frac{\Sigma wz}{\Sigma w} \end{aligned}\right\} \qquad (80)$$

72. Center of mass. If we carry through the same mathematical steps as in the preceding paragraph, but use the masses, $m_1, m_2, \cdots m_n$, of the particles instead of their weights, we obtain the following equations, which are said to define the center of mass.

Defining equations for center of mass:

$$\left.\begin{aligned} \bar{x} &\equiv \frac{\Sigma mx}{\Sigma m} \\ \bar{y} &\equiv \frac{\Sigma my}{\Sigma m} \\ \bar{z} &\equiv \frac{\Sigma mz}{\Sigma m} \end{aligned}\right\} \tag{81}$$

It will be noted that in the determination of these coordinates of the center of mass, the products m_1x_1, m_3y_3, m_nz_n, etc., do not represent torques about an axis as did the products w_1x_1, w_3y_3, w_nz_n, etc., of the preceding section. They are just the products of the elementary masses by their distances from a plane. Accordingly, $\bar{x}$, $\bar{y}$, $\bar{z}$, as given in Eq. (81), do not represent a point through which the line of action of the resultant force acts, but they locate a point whose distance from the plane in question is the **average distance** from that plane of the matter which makes up the body. This is analogous to the "center of population" of a country.

Center of mass is a point which has the important property that, as far as translation is concerned, the entire mass of the body may be considered to be concentrated at the center of mass. Hence, when a number of forces act upon an extended rigid body, the motion of its center of mass is exactly the same as if its entire mass were concentrated at its center of mass and all the forces, unchanged in magnitude and direction, acted at that point. The rotation of the body, however, would not in general be the same as before the change. (See Sec. 148).

Thus when a cat, say, is thrown into the air—or when a diver jumps from a springboard—his center of mass serenely pursues its parabolic path (Sec. 34), regardless of how he may twist and turn; and his linear acceleration and velocity at any instant are exactly the same as if his entire mass were concentrated at his center of mass.

When an unconstrained body is struck a random blow, the

resulting motion will consist of translation of the center of mass in the direction of the force and rotation about the center of mass.

73. When center of gravity and center of mass are the same point. Looking back at Eq. (c) of Sec. 71, the x-coordinate of the **center of gravity** is

$$\bar{x} = \frac{\Sigma wx}{\Sigma w} = \frac{w_1x_1 + w_2x_2 + \cdots + w_nx_n}{w_1 + w_2 + \cdots + w_n}. \tag{a}$$

The weights (w's) being forces, we may express them in absolute units by Eq. (56):

$$w_1 = m_1g_1; \quad w_2 = m_2g_2; \cdots w_n = m_ng_n. \tag{b}$$

Substituting these values in Eq. (a),

$$\bar{x} = \frac{m_1g_1x_1 + m_2g_2x_2 + \cdots + m_ng_nx_n}{m_1g_1 + m_2g_2 + \cdots + m_ng_n}. \tag{c}$$

If the body is large, such as a mountain or a continent, the values of the acceleration of gravity g at its different particles will not all be the same (e.g., g is less at the top of a mountain than at its base), and hence the g's cannot be factored out.

On the other hand, if the body or system of bodies whose center of gravity is being found is so small (e.g., the Great Pyramid) that the values of g at all of its particles are sensibly the same, we may call them all g and factor Eq. (c) so that it becomes:

$$\bar{x} = \frac{\cancel{g}(m_1x_1 + m_2x_2 + \cdots + m_nx_n)}{\cancel{g}(m_1 + m_2 + \cdots + m_n)} = \frac{\Sigma mx}{\Sigma m}. \tag{d}$$

But $\dfrac{\Sigma mx}{\Sigma m}$ is the x-coordinate of the **center of mass**, by Eq. (81).

The same would be true for the y- and z-coordinates, therefore: **The center of gravity is the same point as the center of mass when the body or system of bodies is so small that the acceleration due to gravity is the same at all of its particles.**

As this condition is true for the great majority of the bodies with which we deal, the center of gravity and the center of mass may usually be considered the same point.

74. Location of various centers of mass. Since the masses of bodies made from material of uniform thickness and density, such as sheets of thin metal, are proportional to their areas, we often

speak of the center of mass or the center of gravity of an area in a metaphorical way.

It is shown by means of calculus that the centers of mass of the following bodies are as stated:

1. The center of mass of a figure which is symmetrical with respect to a point is at that point.
2. The center of mass of a parallelogram is at the intersection of the diagonals.
3. The center of mass of a triangle is at the intersection of the medians, which lies one-third of the distance from the middle of any side to the opposite vertex.
4. The center of mass of a pyramid is one-fourth of the distance from the center of mass of the base to the vertex of the pyramid.

75. Computation of center of mass. Center of mass plays an important role in engineering, in which it is usually called the "center of gravity." Its position may easily be calculated if the body can be readily divided up into parts having known centers of mass. The following solved problem illustrates this and shows a convenient method for general use.

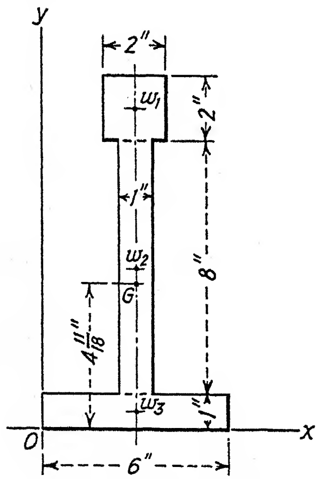


Fig. 72. Deck Beam

Solved Problem

It is required to calculate the location of the "center of gravity" of the cross section of a T-bar, shown in Fig. 72.

If we think of this figure as made of thin metal of uniform thickness and density, the mass of any portion (and its weight) will be proportional to its area. Hence we may use areas in our computation instead of masses or weights.

First divide up the total area into small areas of such form that the position of the center of gravity of each piece is known. It is immaterial just how this is done; but for the given section, division into three rectangles as shown is the most convenient method.

The weights of these pieces, w_1, w_2, w_3, may be considered to act at their respective centers of gravity. Their coordinates and other data for the computation are most conveniently collected into the following table.

Part No.	w (Area)	x	y	wx	wy
1	4	3	10	12	40
2	8	3	5	24	40
3	6	3	0.5	18	3
Σ	18	——	——	54	83

Then, taking the proper values from the table, we have:

$$\bar{x} = \frac{\Sigma wx}{\Sigma w} = \frac{54}{18} = 3 \text{ in.}$$

$$\bar{y} = \frac{\Sigma wy}{\Sigma w} = \frac{83}{18} = 4\tfrac{11}{18} \text{ in.}$$

which are the required coordinates of the center of gravity.

Precisely the same results will be obtained by considering the weights as a system of parallel forces, and solving for the line of action of their resultant. The tabulation, however, has the advantage of orderliness and consequently seems simpler.

PROBLEMS

1. Forces of 5, 3, 8, and 6 lb act simultaneously north, east, south, and west, respectively, at the same point on a body. Find their equilibrant and resultant, graphically and analytically.

2. Forces of 3, 2, 6, and 4 lb act south, east, north, and west, respectively, at the same point and simultaneously. Find their equilibrant and resultant, both graphically and analytically.

3. Forces of 2, 4, 6, and 8 gm act north, south, east, and west, respectively, at the same point on a body. Find their resultant and equilibrant, graphically and analytically.

4. A square $ABCD$ has sides 3 in. long. Forces of 3, 2, 4, and 3 gm act along the sides AB, BC, CD, and AD, respectively. Find their equilibrant and resultant graphically.

5. An equilateral triangle ABC has its sides 2.5 in. long. Forces of 4, 2, and 3 lb act along the sides AB, CB, and AC, respectively. Find their resultant graphically.

6. The cable of an aerial conveyor is 180 ft long and is caused to sag 30 ft at its mid-point by a bucket weighing 400 lb. What is the tension in the cable? What would the tension be if the sag were 15 ft?

7. A loaded bucket weighing 500 lb is at the center of a cable 200 ft long which sags 25 ft at its center. Find the tension in the cable. What would be the tension if it sagged only 10 ft?

8. The ends of a wire 105 ft long are fastened at two points A and B 75 ft apart on a horizontal truss. A weight of 100 lb is attached to the wire at a point 60 ft from B. Find the tension in each section of the wire.

9. A rope 14 ft long is attached to a horizontal beam at two points 10 ft apart. If a weight of 120 lb is attached to the rope at a point 6 ft from one end, what will be the tension in each segment of the rope?

10. A balloon capable of lifting a weight of 640 lb is held against the wind by means of a rope which makes an angle of 60° with the horizontal. Find the tension in the rope and the force of the wind on the balloon.

11. A man weighing 150 lb sits in a hammock supported by two posts. If the ropes make angles of 30° and 45°, respectively, with the horizontal, what is the tension in each rope?

12. A painter weighing 160 lb sits 4 ft from one end of a 16-ft board which is suspended by ropes at its ends. If the board weighs 80 lb, what is the tension in each rope?

13. A 16-ft board is used to make a seesaw. Where must the fulcrum be placed in order that a 90-lb boy may balance one weighing 70 lb?

14. A highway bridge of uniform construction weighs 90 tons. When a 10-ton truck is two-thirds of the way across it, what is the force on each pier?

15. A meter-stick is acted upon by a force of 85 gm upward at 50 cm, 40 gm and 25 gm downward at 10 cm and 90 cm, respectively. Find the equilibrant and the resultant if the weight of the meter-stick is 150 gm.

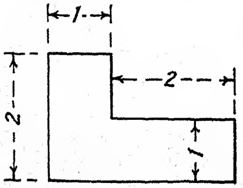
16. A bar 12 ft long, whose weight may be neglected, is held in equilibrium by four forces at points A, B, C, and D. At its ends (A and D), forces of 30 lb and 80 lb, respectively, act downward. Find the forces at B, 2 ft from A; and at C, 4 ft from B.

17. A beam weighing 180 lb is to be carried by three men. Two lift by means of a crossbar and the third lifts at the rear end. Where must the crossbar be placed so that each may bear one-third of the weight of the beam?

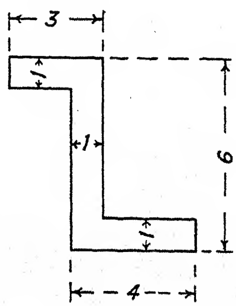
18. A board 16 ft long and weighing 90 lb is suspended by ropes at its ends. Two painters weighing 150 lb and 125 lb stand at 4 ft and at 2 ft, respectively, from the ends. What is the tension in each rope?

19. Find the center of mass (centroid) of this cross section:

20. Find the centroid of this figure:



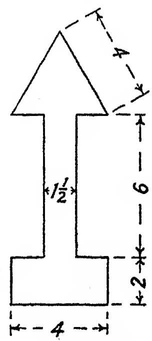
Prob. 19



Prob. 20

21. Find the center of mass (centroid) of this cross section:

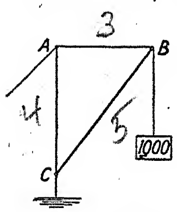
22. A horizontal rod BC, 8 ft long, is attached to a wall at B by a pin-joint and supported at C by a rope, the other end of which is fastened to the wall at A, vertically above B. If the rope makes an angle of 30° with the horizontal, find the tension in AC and the compression in BC when a weight of 50 lb is hung at C.



PROB. 21

23. A horizontal shelf BC, 9 ft long, is hinged to a wall at B and is supported by a rope attached to the wall at A, 8 ft vertically above B, and to the shelf at D, 6 ft from B. A load of 60 lb is suspended at C. Find the stress in the rope and the reaction at B.

24. A crane, as shown, supports a load of 1000 lb. The distances AB, BC, and AC are 12, 20, and 16 ft, respectively. Find the tension in AB and the compression in CB.



PROB. 24

25. A horizontal platform 4 ft wide is hinged at one end to a wall and held at the other end by a rope fastened to the wall 6 ft vertically above the hinge. The platform weighs 200 lb. What is the tension in the rope and the reaction at the hinge?

26. A steel crane consists of a horizontal beam AB 12 ft long, attached to a wall at A and supported at C, 8 ft from A, by a 10-ft strut CD, which is attached to the wall at D vertically below A. A load of 500 lb is suspended from B. If all the joints are pin-joints, find the force in CD and the reaction at A.

27. A 30-ft ladder weighing 150 lb stands on the ground at B and leans against a vertical wall at A. The distance of B from the foot of the wall is 9 ft. Neglecting friction at A, what are the reactions at A and at B?

28. A ladder AB, 20 ft long and weighing 110 lb, rests on the ground at a point B 6 ft from a vertical wall AC. Neglecting the friction at A, find the reactions at A and B (1) graphically, (2) geometrically, and (3) analytically.

29. A ladder 24 ft long rests on the ground at B and against a vertical wall at A. If the weight of the ladder is 120 lb and the horizontal distance of B from the wall is 8 ft, find the reactions on the ladder at A and at B, neglecting friction at A.

30. A bar 16 in. long is acted upon by four forces, as follows: 6 lb upward at 0 in.; 3 lb downward at 6 in.; 5 lb downward at 14 in.; and 2 lb upward at 16 in. Find the resultant.

CHAPTER VI

WORK AND ENERGY

76. Work W is defined as the product of a force F by the distance s through which its point of application moves and the cosine of the angle ϕ between the directions of F and s. In symbols,

$$W \equiv Fs \cos \phi. \tag{82}$$

Thus, in Fig. 73, if the force F acts on the body B and moves it the distance s, the angle between them being ϕ, then, in the scien-

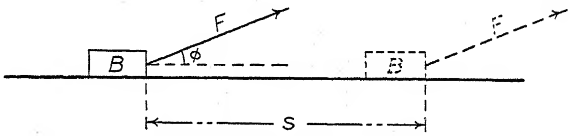


FIG. 73. Force Does Work

tific meaning of the term, the work done by F upon B is given by Eq. (82).

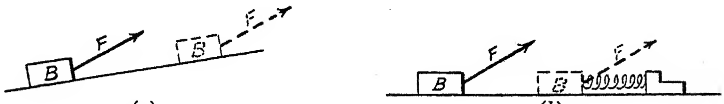
This definition of work is somewhat different from the usual idea of it as held by the non-technical worker. A person ordinarily feels that he is working if he holds up a bucket of water, because he experiences some muscular fatigue. But in the scientific sense, he does no work unless he raises (or lowers) the bucket against (or with) the force of gravity. Similarly, a column exerts force to support a ceiling, but it does no work since it acts through no distance.

When $\cos \phi$ is $+$, the work is said to be done *by the force;* when $\cos \phi$ is $-$, the work is said to be done *against the force.*

The work done by F (Fig. 73) may all be utilized in overcoming the resistance which the horizontal surface offers to B's motion over it. If F is larger than is necessary to overcome that resistance, B will gain speed in accord with Newton's second law. If the body is moved up an incline (Fig. 74a), or is opposed by a spring (Fig. 74b), some of the excess force will do work in pushing the body up the incline, or in compressing the spring.

In general, work is utilized in three ways:

1. In overcoming resistance, like friction.
2. In giving a body speed.
3. In changing the position or condition of a body.



(a) (b)

FIG. 74. Work Stored Up as Energy

The work required to overcome resistance (friction) is usually wasted as heat; but the work done in giving a body speed or in changing its position or condition is stored up in the system as energy and may be recovered, at least in part.

77. Work in rotation.

Special case. To extend the above definition of work to rotation, let the body B of Fig. 75 be rotated through an angle θ by a constant force F unwinding a cord, say, from the cylinder, and therefore acting with the constant lever arm r.

Then the work done by F is, by Eq. (82),

$$W = Fs \cos 0°$$

where s is the length of cord unwound and therefore the distance through which F acts, and being in the same direction as F, $\phi = 0°$.

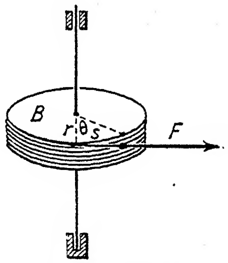


FIG. 75. Work in Rotation

Therefore since $\cos 0° = 1$.

$$W = Fs$$

But by Eq. (2)

$$s = r\theta$$

so that

$$W = Fr\theta$$

also by Eq. (60)

$$Fr = \mathfrak{T}.$$

Therefore

$$W = \mathfrak{T}\theta \qquad (83)$$

where θ must be in radians.

Stating this special case in words, **when the axes of torque and rotation are the same, the work equals the torque multiplied by**

the angle through which the body turns. This special case is the one generally encountered in practice, but it does not show the exact analogy with work of translation as well as does the following case.

General case. In the above special case, F was perpendicular to the axis of rotation; hence F and s were in the same straight line tangent to the cylinder, and $\theta = 0°$.

In the general case, the axis of torque LN (Fig. 76), which is perpendicular to the plane of F and r, does not coincide with the axis of rotation JK, but makes with it an angle ϕ. That is, ϕ is the angle between the directions of the torque vector and the angular displacement vector.

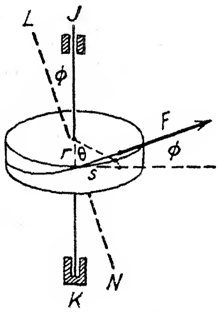


FIG. 76. Work in Rotation (General Case)

Under these circumstances,

$$W = Fs \cos \phi = Fr\theta \cos \phi = \mathfrak{T}\theta \cos \phi. \qquad (84)$$

This equation, which is the exact analogue for rotation of Eq. (82), states that in general the work done in rotation equals the product of the torque $\mathfrak{T}$, the angular displacement θ in radians, and the cosine of the angle ϕ between the directions of the axes of torque and rotation.

Just as in the case of translation, the work done in rotation is employed in overcoming resistance, in giving the body rotational speed, and in changing its position or condition.

78. Energy. As was mentioned in Sec. 76, the work necessary to overcome frictional resistance is transformed into heat and is usually wasted. But the work which is required to elevate a body or to compress a spring may be at least partially recovered; because the body on sliding down the incline, or the spring on returning to its original length, would be capable of doing some work. Also, a body having speed is capable of doing work when it collides with another body. Hence we make the definition: **Energy is the ability to do work.**

It is convenient to say further that:

Potential energy is the energy that a body has by virtue of its position or condition.

Kinetic energy is the energy that a body has by virtue of its motion.

We often speak of the energy of a "body," but this is inaccurate in spite of long established usage. As a matter of fact, energy is always stored in a system of bodies. By **system of bodies** we mean a group of bodies arranged for some purpose; for brevity it is frequently referred to merely as "a system."

For example, when water is pumped into an elevated tank we say that "energy is stored in the water"; but the water is able to do work only because it has been moved farther away from the center of gravity of the earth. Hence the energy is really stored in the system consisting of the water and the earth. In the case of the chemical energy "stored in coal," the energy is due to the work done in the separation of the carbon, hydrogen, sulphur, etc., from the oxygen; so that the system really consists of the fuel and the oxygen.

However energy be manifested—in heat, in light, in a radio signal, or in a tornado—it is believed to be one and the same thing, and to be indestructible.

79. Mathematical expressions for energy. Since energy is the ability to do work, the logical way to measure energy is by the amount of work the system can do when giving up its energy. In order to avoid obscuring the main facts at first by too many details, it is customary to neglect work wasted in overcoming frictional resistance, which in many cases may be a small fraction of the total work. We accordingly begin with the **ideal case** in which the energy stored in a system, when in a given condition, equals the work necessary to get it into that condition.

Potential energy. In the ideal case, potential energy (P.E.) is given by the expression:

$$\left.\begin{aligned} \text{P.E.} &\equiv W \equiv Fs\cos\phi \\ \text{or}\qquad \text{P.E.} &\equiv W \equiv \mathfrak{L}\theta\cos\phi \end{aligned}\right\} \tag{85}$$

where W is the work done upon the body to get it into the position or condition in which it has potential energy.

Kinetic energy of translation. Similarly, in the ideal case, the kinetic energy (K.E.) of a body equals the work that must be

done upon it to bring the body up to the speed it has at the time we are considering it. Therefore,

$$\text{K.E.} \equiv W \equiv Fs \cos \phi. \tag{a}$$

When we see an automobile speeding along, it is obviously impossible to determine at that time how much force acted upon it and through what distance it acted in getting the body up to that speed. Moreover, since kinetic energy is the energy a body has by virtue of its speed, it is most natural to express kinetic energy in terms of *speed*. This is easily accomplished.

From Eq. (53), $F = Ma$ (provided absolute or B.E. units are used)

and from Eq. (26), $s = \frac{1}{2}at^2$.

If we have the push F parallel to the displacement s, which is the simplest way, then $\phi = 0$; $\cos \phi = 1$.

Substituting these values in Eq. (a), above, we have:

$$\text{K.E.} = (Ma)(\tfrac{1}{2}at^2) = \tfrac{1}{2}Ma^2t^2 = \tfrac{1}{2}M(at)^2. \tag{b}$$

But by Eq. (25), $v = at$.

Substituting this for at in Eq. (b),

$$\text{Kinetic energy of translation of a body} = \tfrac{1}{2}Mv^2 \tag{86}$$

provided that absolute or B.E. units are used.

Kinetic energy of rotation. An algebraic expression for the kinetic energy of a rigid body due to its rotational motion is readily derived from the foregoing expressions for kinetic energy of translation. Consider a rigid body B (Fig. 77), having an angular speed of ω rdn/sec about an axis perpendicular to the paper at A.

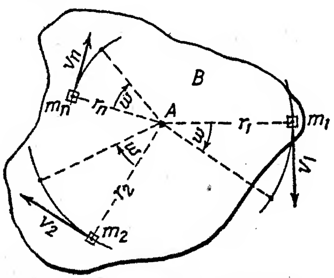


FIG. 77. Energy of Rotation

Let the body be divided up into elementary masses $m_1, m_2, \ldots m_n$,* whose distances from the axis are $r_1, r_2, \ldots r_n$, respectively. Each of these elementary masses will describe a circle about A. Call their linear speeds around these circles $v_1, v_2, \ldots v_n$, respectively.

* Just how small these masses must be depends upon circumstances. If necessary, they may be given the limiting value dm, as in calculus.

Then, from Eq. (86),

$$\begin{aligned}
(\text{K.E.})_1 \text{ of } m_1 &= \tfrac{1}{2}m_1v_1^2 \\
(\text{K.E.})_2 \text{ of } m_2 &= \tfrac{1}{2}m_2v_2^2 \\
\vdots \qquad \vdots \quad &\qquad \vdots \\
(\text{K.E.})_n \text{ of } m_n &= \tfrac{1}{2}m_nv_n^2.
\end{aligned}$$

Since kinetic energy is a scalar quantity, the total kinetic energy of the body due to its rotation is obtained by adding these elementary quantities. Therefore,

$$\text{K.E. of body} = \tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2 + \cdots + \tfrac{1}{2}m_nv_n^2.$$

The body being rigid, all of its particles will have the same angular speed ω about the axis; hence by Eq. (35),

$$v_1 = r_1\omega;\ v_2 = r_2\omega;\ \cdots v_n = r_n\omega.$$

Substituting these values above,

$$\begin{aligned}
\text{K.E. of body} &= \tfrac{1}{2}m_1r_1^2\omega^2 + \tfrac{1}{2}m_2r_2^2\omega^2 + \cdots + \tfrac{1}{2}m_nr_n^2\omega^2 \\
&= \tfrac{1}{2}\omega^2(m_1r_1^2 + m_2r_2^2 + \cdots + m_nr_n^2).
\end{aligned}$$

But by Eqs. (63) and (64),

$$(m_1r_1^2 + m_2r_2^2 + \cdots + m_nr_n^2) \equiv \Sigma mr^2 \equiv I.$$

Therefore,

$$\text{Kinetic energy of rotation of a body} = \tfrac{1}{2}I\omega^2 \tag{87}$$

provided that absolute or British engineering units are used.

The total kinetic energy of a body having both translation and rotation at the same time is obviously:

$$\text{K.E.} = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\omega^2$$

since energy is a scalar quantity.

80. Conservation of energy. In the expressions for potential energy of Eq. (85), the displacements s and θ are measured from, or relative to, some point or line considered arbitrarily as fixed; and the linear speed v of Eq. (86) and the angular speed ω of Eq. (87) are likewise relative to some convenient frame of reference considered as at rest. Hence all these expressions give the energy of one body as measured relative to another body, or of a system in one configuration as measured relative to another configuration of that system.

Thus, the water in a tank on a roof has a certain amount of potential energy relative to the top floor of the building, and still more potential energy relative to each of the lower floors. Similarly, a person walking through a moving train has a certain amount of kinetic energy relative to the train and yet more, or less, relative to the earth, depending upon whether he is walking in the same or the opposite direction to the motion of the train, and upon their relative speeds.

From these facts it is evident that the measurement of energy is always relative.

However, when one system loses energy some other system always gains an exactly equal amount, as is verified by experiment. Hence we are led to believe that energy is an *entity* of nature which is conserved; and we state the **law of conservation of energy: The total amount of energy in the universe remains constant.** In other words, energy can neither be created nor destroyed.

This law, first enunciated by Robert Mayer in 1842, has formed with the law of conservation of matter (Sec. 43) the foundation on which the whole structure of physical science has been built. Since 1932, however, research seems to have established definitely the fact that in small quantities energy can be converted into matter and matter into energy (see Sec. 657). The two conservation laws therefore become one.

The amounts of energy as we measure them are relative, because our system of measuring is limited to such measurement. Energy itself appears to be an entity having an absolute existence quite independent of our ability to determine its amount.

An important law of energy which we shall frequently find exemplified is: **The potential energy of a system tends to become a minimum.** Thus, stones tend to roll down hill, springs to unwind, batteries to discharge.

81. Units of work and energy. The units of work and energy are derived by writing into Eq. (85) the units of force and distance which we already have.

$$W \text{ dyne-cm} = (F \text{ dynes})\ (s \text{ cm})\ (\cos \phi).$$

Cos ϕ being a pure number, no unit is shown for it.

By carrying out this process for the several systems, we obtain the following units.

Absolute systems:

The $\left\{\begin{array}{l}\text{dyne-centimeter}\\ \text{foot-poundal}\end{array}\right\}$ is defined as the work done when a force of $\left\{\begin{array}{l}\text{1 dyne}\\ \text{1 poundal}\end{array}\right\}$ acts through a distance of $\left\{\begin{array}{l}\text{1 cm}\\ \text{1 ft}\end{array}\right\}$ in the direction of the force (i.e., $\cos \phi = 1$).

Gravitational systems:

The $\left\{\begin{array}{l}\text{gram-centimeter}\\ \text{foot-pound}\end{array}\right\}$ is defined as the work done when a force of $\left\{\begin{array}{l}\text{1 gram}\\ \text{1 pound}\end{array}\right\}$ acts through a distance of $\left\{\begin{array}{l}\text{1 cm}\\ \text{1 ft}\end{array}\right\}$ in the direction of the force.

The dyne-centimeter is called an *erg* (Greek, work). Hence, **an erg is the amount of work done when a force of 1 dyne acts through a distance of 1 centimeter in the direction of the force.**

The erg is a very small amount of work. A postage stamp weighs about 50 dynes, so the work required to lift it 1 cm is about 50 ergs. For practical purposes a much larger unit is desirable. Consequently, we have the **joule,** which is **defined as 10 million ergs.**

$$1 \text{ joule} \equiv 10^7 \text{ ergs.} \tag{88}$$

The reason for choosing this particular power of 10 in the definition is that it makes the expression for electrical work the same in the practical system of units as in the electrostatic and electromagnetic systems. It has the accidental advantage that it makes the joule a unit of the same order of magnitude as the foot-pound (1 joule = 550/746 ft-lb). Had 10^8 or 10^6 been used, the discrepancy between these two practical units would have been very much greater.

Solved Problem

1. Given a body of mass M at a height h above the earth. Required its potential energy with respect to the earth; and its speed on reaching the earth, if allowed to fall freely.

To lift a body against the force of gravity will require a force F equal and opposite to the weight of the body, plus a small excess at first to give initial acceleration. But it is customary to neglect this slight excess and to say that the necessary force equals the weight of the body.

In gravitational units, the weight of a body is *numerically* equal to its mass (Sec. 49). Hence we have the values:

$$F = M \text{ gm of force (say)}$$
$$s = h \text{ cm}$$
$$\cos \phi = 1 \quad \text{since } F \text{ and } h \text{ are parallel.}$$

Substituting in Eq. (85),

$$\text{P.E.} = Mh \text{ gm-cm}$$
$$= Mgh \text{ dyne-cm, or ergs.}$$

A similar argument will apply to the British units, hence we may say in general for an elevated body:

$$\left.\begin{aligned}\text{P.E.} &= Mh \text{ gravitational units} \\ &= Mgh \text{ absolute units.}\end{aligned}\right\} \tag{89}$$

Let the body fall freely back to the earth from rest at the height h, and let v be the speed with which it reaches the earth.

Then its kinetic energy when it strikes the earth will be:

$$\text{K.E.} = \tfrac{1}{2}Mv^2 \text{ absolute units}$$

and its potential energy will be zero with respect to the earth, since it is no longer elevated.

By the conservation law, energy can be neither created nor destroyed; consequently, the body must reach the ground with as much energy as it had when it started. (We neglect for the present the energy wasted in overcoming the friction of the air.) Therefore,

$$\tfrac{1}{2}Mv^2 = Mgh$$

or

$$v = \sqrt{2gh} \tag{90}$$

which is the speed with which the body would strike the ground. This same result has already been found by another method, in Eq. (28).

2. A plane is 20 meters long and inclined 15° to the horizontal. How long will it take a circular disk of mass M gm and radius r cm to roll down the plane without sliding from rest at the top, under the action of its own weight? (Fig. 78.)

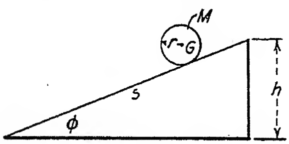


FIG. 78

Analysis: When the body is at rest at the top of the plane, its energy is entirely potential; when it reaches the bottom, its energy is entirely kinetic. By the **principle of conservation of energy** these two amounts of energy must be equal, since no work is done by the body during its descent.

Known:

$$s = 20 \text{ m} = 2000 \text{ cm} \qquad r = r \text{ cm}$$
$$\phi = 15°; \sin 15° = 0.259 \qquad v_1 = 0$$
$$M = M \text{ gm.}$$

Solution:

At top, Potential energy $= Mgh$ ergs by Eq. (89).

At bottom,

Translational kinetic energy $= \frac{1}{2}Mv^2$ ergs by Eq. (86)

Rotational kinetic energy $= \frac{1}{2}I\omega^2$ ergs by Eq. (87).

By the law of conservation of energy, as above,

$$Mgh = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\omega^2.$$

But by Sec. 58,

$$I = M\frac{r^2}{2}$$

and by Eq. (34),

$$\omega = \frac{v}{r}.$$

Hence,

$$Mgh = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}\left(M\frac{r^2}{2}\right)\left(\frac{v^2}{r^2}\right)$$

$$gh = \tfrac{3}{4}v^2$$

whence,

$$v = \sqrt{\frac{4gh}{3}}. \qquad \text{(a)}$$

From the figure,

$$\begin{aligned} h &= s \sin \phi \\ &= 2000 \times 0.259 = 518 \text{ cm}. \end{aligned}$$

Substituting in Eq. (a),

$$v = \sqrt{\frac{4 \times 981 \times 518}{3}} = 823 \text{ cm/sec}.$$

This v is the final speed at the bottom. Since the motion is uniform, the average velocity is:

$$\bar{v} = \frac{v + 0}{2} = \frac{823}{2} = 411.5 \text{ cm/sec}$$

and the time t required to roll down is

$$t = \frac{2000}{411.5} = 4.86 \text{ sec}.$$

By comparing this with Problem 2, Sec. 28, it will be noted that the time for a body to *slide* down the same plane is only 3.97 sec.

82. Power. It will have been observed that no mention of time is made in the definition of work. However, if one winding engine hauls up in 2 minutes a car of coal which another engine requires 5 minutes to raise, the former is in that respect the better engine. To cover cases of this kind we define power.

Power, P, is the time rate of doing work. Algebraically,

$$\text{Average power, } P \equiv \frac{W}{t}. \tag{91}$$

$$\text{Instantaneous power, } P \equiv \frac{dW}{dt}. \tag{92}$$

The term **activity** is sometimes used as a synonym for power.

83. Units of power are obtained by dividing the units of work by the desired unit of time. Thus the **foot-pound per second** and the **erg per second** are the two basic units of power. Both, however, are rather small for general business, so multiples of them are taken as **practical units of power:**

$$1 \text{ horsepower (hp)} \equiv 550 \frac{\text{ft-lb}}{\text{sec}}$$

$$1 \text{ watt} \equiv 1 \frac{\text{joule}}{\text{sec}} = 10^7 \frac{\text{ergs}}{\text{sec}}$$

$$1 \text{ kilowatt (kw)} \equiv 1000 \text{ watts.}$$

From these definitions it may easily be shown that

$$1 \text{ hp} \equiv 746 \text{ watts.}$$

For an 8-hr working day, a manpower is about 1/10 hp.

84. Dimensional formulas for work, energy, and power. By putting the proper fundamental units in their respective defining equations, the dimensional formulas for work, energy, and power may readily be obtained.

Work: From Eq. (82), $W \equiv Fs \cos \phi$.

From Eq. (53a) $[F] = [MLT^{-2}]$

and by Sec. 10,

$$[\cos \phi] = [1].$$

Hence,

$$\begin{aligned}[W] &= [MLT^{-2}]\,[L]\,[1] \\ &= [ML^2T^{-2}].\end{aligned} \tag{93}$$

Kinetic Energy: From Eq. (86), K.E. $= \frac{1}{2}Mv^2$.

The dimensional formula for $[v]$ is $[LT^{-1}]$ by Sec. 10.

Hence,

$$\begin{aligned}[\text{K.E.}] &= [1]\,[M]\,[LT^{-1}]^2 \\ &= [ML^2T^{-2}].\end{aligned} \tag{94}$$

It will be noted that the dimensional formulas for work and kinetic energy are the same. This is as it should be, since work and energy are different aspects of the same thing.

Power: By its definition,

$$P \equiv \frac{W}{t}$$

$$= \frac{[ML^2T^{-2}]}{[T]}$$

$$= [ML^2T^{-3}]. \tag{95}$$

It should be carefully noted that while quantities of the same physical nature invariably have the same dimensional formulas, the converse is not true. For example, work and torque (Sec. 54) both have the dimensional formula $[ML^2T^{-2}]$; but physically they are quite different, work being a scalar and torque a vector.

85. Friction is defined as the resisting force, or resistance, which one body offers to the sliding or rolling of another body over it.

Sliding friction is due to the interlocking of the asperities of two surfaces, as in Fig. 79. Even highly polished surfaces show these irregularities when greatly magnified. If the friction is sufficient to prevent relative motion of the bodies, it is said to be **static friction;** but if the bodies move with respect to each other, it is called **kinetic friction.**

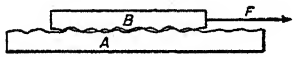


FIG. 79. Sliding Friction

Laws of sliding friction between solids. These approximate laws are from the work of the French General A. T. Morin (1831) and may be summarized as follows:

1. Friction between solids depends upon:
   (a) Nature of the substances; e.g., iron, leather, wood, etc.
   (b) Condition of the surfaces; e.g., how smooth, whether lubricated, etc.
   (c) The normal force pressing the surfaces together, and is proportional to that force.
2. Friction between solids is independent of:
   (a) The area of the surfaces in contact.
   (b) The speed, except at starting.
3. Static friction is greater than kinetic friction.

4. Kinetic friction is greater than rolling friction, other things being equal.

Friction is greatest when motion is impending; i.e., when one body is on the point of starting to move with respect to the other. Then, writing law 1(c) above in algebraic symbols,

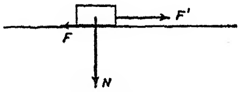


FIG. 80

$$F \propto N$$
$$F = \mu N$$
$$\mu \equiv \frac{F}{N}. \tag{96}$$

The coefficient of static friction μ is defined as the ratio of the maximum force of friction F to the total normal force N which presses the surfaces together.

As is shown in Fig. 80, the friction F is equal and opposed to the force F', which is just sufficient to start the body B in motion. The force F'' necessary to keep B in uniform motion with respect to A is slightly less than F'. The ratio of F'' to N is called the **coefficient of kinetic friction.**

86. Angle of repose. A simple and convenient method of finding the coefficients of sliding friction is the following:

Let the surface OX (Fig. 81) be tilted up gradually until the body B just begins to slide down the incline. If the weight W of the body is resolved into its components F' parallel and N normal to the incline, it is evident that F' alone overcomes the friction F, while N merely presses the bodies together.

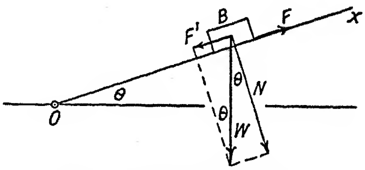


FIG. 81. Angle of Repose

Just before sliding begins, there is no acceleration. Hence B is in equilibrium, and by Eq. (75),

$$\Sigma F_x = 0.$$

That is,

$$F - F' = 0$$

or

$$F = F'.$$

Hence, by definition (Eq. 96),

$$\mu \equiv \frac{F}{N} = \frac{F'}{N}.$$

But from Fig. 81,

$$\frac{F'}{N} = \tan \theta.$$

Therefore,

$$\mu = \tan \theta \tag{97}$$

where μ is the coefficient of **static friction.**

The **angle of repose** θ is the maximum angle of inclination that a plane may have before a body on the plane begins to slide under the action of its weight alone.

After sliding begins, the angle may be decreased slightly to a value θ'', so that the body will slide without acceleration. The angle θ'' is called the limiting angle of **kinetic friction,** and the the coefficient of **kinetic friction,** is

$$\mu'' = \tan \theta''. \tag{98}$$

87. Rolling friction. Rolling friction is the resistance arising from the fact that the body rolled upon is slightly indented and the rolling body slightly flattened where they are in contact. A short time is required for the bodies to recover from these deformations, which are usually temporary. The result is that the rolling body B is constantly climbing a slight grade, as is shown—greatly exaggerated—by the line JK in Fig. 82.

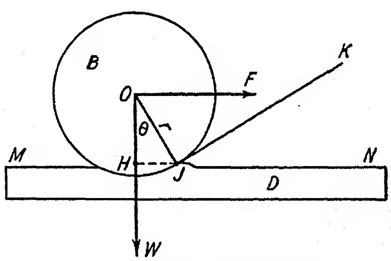


FIG. 82. Rolling Friction

At any instant, the reaction of the surface MN passes through some point J; and if the motion is uniform, the body B is in equilibrium, so that this reaction must pass through the center O also. Since the body is in equilibrium, we have, by Sec. 64:

$$\Sigma \mathfrak{J} = 0.$$

Taking the axis of torques at J, and calling $HJ \equiv k$, we have:

$$Wk - Fr\cos\theta = 0$$

$$F = k\frac{W}{r\cos\theta} \tag{99}$$

where F is the force necessary to overcome the rolling friction and W is the normal force pressing the bodies together.

The constant k is called the **coefficient of rolling friction** and is seen to have the dimensional formula $[L]$. It is practically independent of the speed of rolling and of the temperature, but depends upon the materials of which the bodies are made. Some approximate values of k are as follows:*

Steel on steel	0.01 in.
Hardwood on hardwood	0.02 in.
Pneumatic tires on good road	0.022 in.

The laws of rolling friction are not yet well known. However, the force F required to overcome rolling friction is seen by Eq. (99) to vary inversely as the radius of the roller. The smaller the depression, the greater is $\cos\theta$; hence, the more rigid the surfaces, the less the rolling friction.

88. Efficiency. Because the work done in overcoming friction (and similar resistances) is usually wasted as heat, the useful work that we can get out of a system is always less than the total work (energy) that we put into, or do upon, the system.

Total work = Useful work + Work wasted in overcoming friction.

Efficiency is the ratio of the useful work got out of a system to the total work put into, or done upon, the system.

This definition has been conveniently abbreviated by Professor C. R. Mann into the form:

$$\text{Efficiency} \equiv \frac{\text{Work out}}{\text{Work in}}. \tag{100}$$

Efficiency is always a fraction less than unity. It is usually expressed in percentage.

* F. B. Seely and N. E. Ensign, *Analytical Mechanics for Engineers*, 2d ed. (New York, John Wiley & Sons, 1933), p. 125.

Example. A 1-hp motor will do 33,000 ft-lb of work per minute on a pumping system. But it will *not* elevate 330 lb of water per minute into a tank 100 ft high. Its performance would be somewhat as follows.

An ordinarily good pump might raise 198 lb of water into the tank per minute. In that case we should have, per minute:

Total *work* put *in* by motor	33,000 ft-lb
Useful *work* got *out* = 198 × 100 × 1 =	19,800 ft-lb
Work wasted in friction =	13,200 ft-lb.

The efficiency of the pumping system, including the pump and the pipe lines, would then be:

$$\text{Efficiency} \equiv \frac{\text{Work out}}{\text{Work in}} = \frac{19{,}800 \text{ ft-lb}}{33{,}000 \text{ ft-lb}} = 60\%.$$

Obviously, when a system is in continuous operation, the time factor for the work out is the same as the time factor for the work in, so that

$$\text{Efficiency} = \frac{\text{Power out}}{\text{Power in}}. \tag{101}$$

Solved Problem

A bullet having a mass of 25 gm, and a speed of 400 meters per sec, passes through a plank 5 cm thick which offers an average resistance of 350 kg of force. With what speed will the bullet come out?

Analysis: The bullet has a certain amount of kinetic energy when it strikes the board. Part of this energy is utilized in doing the work of overcoming the resistance offered by the board. The remaining energy corresponds to the speed with which the bullet comes out of the board, and from it that speed can be computed.

Known:

Mass of bullet, $M = 25$ gm
Initial speed of bullet, $v_1 = 400$ m/sec $= 40{,}000$ cm/sec
Thickness of plank, $s = 5$ cm
Average resistance of plank, $F = 350$ kg $= 350{,}000$ gm force.

Required:

Final speed, $v_2 = ?$

Solution:

$$\begin{aligned}\text{Initial kinetic energy} &= \tfrac{1}{2}Mv_1^2 && \text{from Eq. (86)}\\ &= \tfrac{1}{2} \times 25 \times (40{,}000)^2\\ &= 2 \times 10^{10} \text{ ergs}\\ &= 2000 \text{ joules.}\end{aligned}$$

$$
\begin{aligned}
\text{Work to penetrate board} &= Fs \cos \phi \qquad \text{from Eq. (82)} \\
&= 350{,}000 \times 5 \times 1 \\
&= 175 \times 10^4 \text{ gm-cm} \\
&= 980 \times 175 \times 10^4 \text{ ergs} \\
&= 1715 \times 10^6 \text{ ergs} = 171.5 \text{ joules.} \\
\text{Final kinetic energy} &= 2000 - 171.5 = 1828.5 \text{ joules.}
\end{aligned}
$$

That is,

$$
\begin{aligned}
\tfrac{1}{2}Mv_2^2 &= 1828.5 \text{ joules} = 18285 \times 10^6 \text{ ergs} \\
\tfrac{1}{2} \times 25 \times v_2^2 &= 18285 \times 10^6 \text{ ergs} \\
v_2 &= 38250 \text{ cm/sec} \\
&= 382.5 \text{ m/sec.}
\end{aligned}
$$

PROBLEMS

1. A man weighing 150 lb sits in a wagon weighing 300 lb and is drawn by a horse along a level road at 3 mph. The tension in each trace is 50 lbf and they make an angle of 4° with the horizontal. How much power is expended by the horse, wagon, and man, respectively?

2. A box whose mass is 50 lb rests on a horizontal platform. If the coefficient of friction is 0.3, how much work is required to slide it 60 ft? What speed will it be given in 20 sec by a horizontal force of 25 lb?

3. Find the work done by a man carrying 100 lb of brick up a ladder 20 ft long making an angle of 30° with the wall.

4. If the coefficient of friction of wood on wood is 0.35, how much work is necessary to drag an oak beam weighing 500 lb up a wooden ramp 40 ft long inclined 20° to the horizontal?

5. A mass of 10 lb falls from rest at a height of 144 ft. Make a table showing its potential, kinetic, and total energies at the beginning of its fall and at the end of each second thereafter.

6. An ice boat weighs 250 lb. If the coefficient of friction of steel on ice is 0.02 and the net wind force is 70 lb, what speed will it acquire in 5 sec?

7. A steam engine makes 250 rpm and has a stroke of 18 in. If the average normal force of the crosshead against the guides is 100 lb and the coefficient of friction is 0.044, how much power is wasted in moving the crosshead?

8. A body is projected along a horizontal plane with a speed of 100 ft/sec. If the coefficient of friction is 0.1, find how far the body will move.

9. A boy weighing 90 lb is 3 ft from the axis of a joy wheel. If the coefficient of friction of cloth on wood is 0.2, at what speed will the boy begin to slip?

10. The rotor of a motor weighs 300 lb. If the coefficient of bearing friction is 0.06, and the shaft 2 in. in diameter, what power (kilowatts) is required to maintain a speed of 3600 rpm at no load?

11. Find the horsepower of an engine which pumps 5 ft³ of water per minute from a mine 1/4 mi deep. The engine has an efficiency of 20%.

12. What must be the power of a winding engine to haul a ton of coal in a mine car which weighs 400 lb up a 30° incline 1/2 mi long in 3 min?

13. If a loaded mine car weighs 2500 lb and the mine slope is 0.5 mi long and slopes 10° from the horizontal, what must be the power of a winding engine to haul up the car in 4 min?

14. The muzzle velocity of a 300-lb shell from a gun weighing 50 tons is 2000 ft/sec. Find the kinetic energy of the shell as it leaves the gun, and the velocity and the kinetic energy of the gun. (Momentum is conserved.)

15. The hammer of a pile driver weighs 500 lb. If it falls from a height of 12 ft and drives a pile 1 in., what is the average resistance offered to the pile by the earth?

16. An auto whose mass is 3000 lb is moving at 60 mph when it collides with a car of the same kind parked with its brakes on. The cars slide 25 ft. What was the average force of the impact?

17. A bullet having a mass of 1/2 oz strikes a board 1 in. thick with a speed of 10,000 ft/sec and emerges with a speed of 5000 ft/sec. What is the average resistance offered by the board?

18. A bullet of 20 gm mass and moving at a speed of 40,000 cm/sec strikes a tree. To what depth will it penetrate if the average resistance is 4×10^8 dynes?

19. A bullet whose mass is 12 gm strikes a board 6 cm thick at a speed of 30,000 cm/sec. If the average resistance is 5×10^8 dynes, with what speed will it emerge?

20. If the coefficient of friction between a slide-valve of an engine and its seat is 0.05, how much power is required to operate a valve weighing 50 lb if the travel is 10 in. and the engine speed 250 rpm?

21. What must be the power of a motor to pump 1000 gal of water per hour into a tank (average height 90 ft), assuming an efficiency of 60%?

22. What is the efficiency of a pumping system driven by a 10-hp motor which elevates 50 ft³ of water per minute to an average height of 100 ft?

23. What is the efficiency of a pumping system if a 10-hp engine elevates 120 ft³ of water into a tank (average height 80 ft) in 3 min?

24. A horse draws a wagon up a hill 440 ft long measured on the slope, and rising vertically 20 ft. If the wagon weighs 1500 lb and the horse exerts a force of 180 lb parallel to the incline, find the total work done by the horse, the work done against gravity, and the efficiency of the system.

25. If the coefficient of friction is 0.05, and the average normal force between a crosshead and its guides is 110 lb, how much power is wasted in moving the crosshead when the engine makes 300 rpm, the stroke being 24 in.?

26. What power is transmitted by a belt running 3600 ft/min if the tension on the slack side is 40 lb and that on the tight side is 90 lb?

27. A belt transmitting 10 hp is running at a speed of 3000 ft/min. If the tension on the slack side is 30 lb force, what is it on the tight side?

28. A centrifuge is to make 1200 rpm when the handle makes 90. If the gear on the shaft of the centrifuge has 12 teeth, how many teeth must there be on the gear on the handle?

29. What torque is required to bring up to a speed of 200 rpm in 40 sec a flywheel whose mass is 1200 lb and radius of gyration 2.5 ft? How many revolu-

tions does it make in coming up to this speed and how much work is done upon it?

30. The rotor of a turbine weighs 1600 lb and has a radius of gyration of 2 ft. Neglecting friction, what torque will be required to bring it up to a speed of 150 rpm in 30 sec? How many revolutions does it make in coming up to this speed and how much work must be done upon it?

CHAPTER VII

MECHANICS OF FLUIDS

89. Mechanics of fluids, or **hydrodynamics,** is divided into *hydrostatics*, which treats of fluids at rest, and *hydraulics*, which treats of fluids in motion.

Fluids were defined in Sec. 5 as substances that flow, and include liquids and gases. In addition to the distinctions there given, it should now be noted that *liquids can be compressed only with great difficulty, but gases may be compressed with great ease.* Thus, water is reduced only 0.00005 in volume by an increase in pressure of 1 atmosphere, whereas air is reduced one-half. Consequently, for most purposes liquids may be considered to be incompressible.

An **ideal,** or **perfect, fluid** is defined as one that would offer no resistance to steady flowing. Such a fluid, if set whirling in a bucket, would continue to rotate forever. No ideal fluid exists: all actual fluids offer resistance to steady flowing—some very little, such as air and gasoline; others very much, such as molasses and pitch.

Actual fluids, if given time, eventually yield and flow under the action of any shearing forces, however small. That is, the modulus of rigidity (Sec. 178) of a fluid is zero.

The resistance which a fluid offers to steady flowing depends upon a property called its viscosity, as will be seen from the following article.

90. Laws of fluid, or viscous, friction. When a fluid flows along a surface, as pitch down a board (Fig. 83), the particles in each layer of the fluid are retarded by those in the next lower layer, which is moving more slowly, on account of the friction between the two layers.

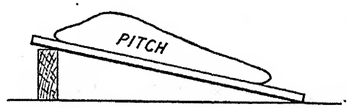


FIG. 83. A Viscous Substance

In order to obtain quantitative relations, consider a plane sur-

face MN (Fig. 84), parallel to a very large plane surface OX, from which it is separated by a layer of some fluid. In general, a film of the fluid will adhere to the lower side of MN and to the upper side of OX, and these films will move with those surfaces respectively. Hence, when MN is moved uniformly and parallel with respect to OX as shown, the force required is that necessary to overcome the friction of these attached films against the intervening layer of fluid.

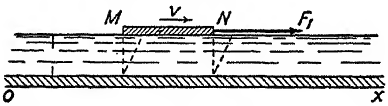


FIG. 84. Fluid Friction

Experiment shows that the force F_1 necessary to give MN a uniform velocity v with respect to OX—i.e., the force required to overcome the fluid friction F—will be in accord with the following approximate laws:

The friction F between two layers of a fluid varies:

1. Directly as the area A of the surface of contact.
2. Directly as the velocity v of one layer with respect to the other, up to a certain critical velocity; and as the square of their relative velocity above the critical value.
3. Inversely as the thickness h of the whole layer of fluid between the restraining surfaces.
4. Independently of the normal force pressing the surfaces together as long as the layer of fluid is not broken through.

Stating these facts algebraically,

$$F \propto \frac{Av}{h}$$

$$F = \eta \frac{Av}{h}. \tag{102}$$

The constant of proportionality η in the above equation is called the **coefficient of viscosity.**

Observation shows, as would be expected, that η is different for different substances. Hence it is customary to add, as another law of fluid friction:

5. Fluid friction varies directly as the viscosity of the fluid.

91. Viscosity. Examination of Eq. (102) reveals that η is the only factor which depends upon the nature of the fluid. Hence it represents a property of the fluid, and we form the qualitative definition:

Viscosity is that property of a fluid that determines its resistance to steady flowing. Quantitatively, viscosity is measured by the coefficient η.

On solving Eq. (102) for η, we obtain the defining equation:

$$\textbf{Coefficient of viscosity, } \eta \equiv \frac{Fh}{Av}. \tag{103}$$

In general, when we speak of the viscosity of a fluid, we mean its coefficient of viscosity.

The viscosity of liquids decreases as the temperature increases. This is a matter of common experience, for all are acquainted with the fact that it is more difficult to crank an automobile when it is cold (viscosity of oil great) than when it is hot (viscosity of oil small).

For gases, on the other hand, the viscosity increases as the temperature increases.

Viscosity is one of the most important properties of lubricating oils. An oil must be selected with sufficient viscosity to maintain a continuous layer between the moving pieces (such as a shaft and its bearing) so that they shall at no time come in contact; but the viscosity must not be so great as to cause unnecessary waste of energy to overcome it.

The work done in overcoming viscosity, like that in overcoming friction between solids, is transformed into heat and wasted.

92. Unit of viscosity: the poise. Imagine that in Fig. 85 the upper plane has an area of 1 cm², is parallel to the lower plane, which is very large, and is separated therefrom by a layer 1 cm thick of some fluid. Suppose also that a force of 1 dyne acting on and parallel to the upper surface is just sufficient to maintain a velocity of 1 cm/sec of that surface with respect to the lower surface. We could then find the viscosity of this fluid by substituting these values in Eq. 103, thus:

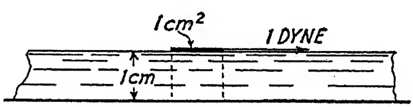


FIG. 85. Poise Defined

$$\begin{aligned}\eta &\equiv \frac{Fh}{Av}\\ &= \frac{1\text{ dyne} \times 1\text{ cm}}{1\text{ cm}^2 \times 1\text{ cm/sec}}\\ &= 1.\end{aligned}$$

That is, the fluid in which these data would be obtained would have *unit viscosity*.

This cgs absolute unit of viscosity is called the *poise:*

$$1 \text{ poise} \equiv \frac{1 \text{ dyne} \times 1 \text{ cm}}{1 \text{ cm}^2 \times 1 \text{ cm/sec}}. \qquad (104)$$

In words the definition would be as follows:

The **poise** is the viscosity of a fluid such that if two parallel planes, the area of one being 1 cm² and the other of very large area, are separated by a layer of the fluid 1 cm thick, a force of 1 dyne acting on and parallel to the small surface will maintain for it a uniform velocity of 1 cm/sec relative to the large surface.

The **centipoise** (≡ 1/100 poise) is the unit of viscosity most generally used.

Specific viscosity is the ratio of the viscosity of the substance to that of water when both are at 0°C.

HYDROSTATICS

93. Pressure P has been defined in Sec. 38 as force F per unit of area A. This gives the average pressure over the area. In algebraic form,

$$\textbf{Average pressure, } P \equiv \frac{F}{A}. \qquad (105)$$

But when the pressure varies from point to point of the surface, the defining equation becomes:

$$\textbf{Pressure at a point, } P \equiv \frac{dF}{dA}. \qquad (106)$$

Units of pressure are obviously obtained from the units of force and area which are used in any given case. Thus, in the systems already mentioned we have:

$$\text{Gravitational}\begin{cases}\text{metric system—gram force/cm}^2\\ \text{British system—pound force/ft}^2\end{cases}$$

$$\text{Absolute}\begin{cases}\text{cgs system—dyne/cm}^2\\ \text{fps system—poundal/ft}^2.\end{cases}$$

By international agreement, special names have been given to two units of pressure:

The **barye** is defined as 1 dyne per cm².

The **bar** is defined as 1,000,000 dynes per cm².

But in the United States it is quite generally the custom to ignore these definitions and to take the unit:

$$1 \textbf{ bar} = 1 \text{ dyne/cm}^2.$$

Great care must be exercised when reading scientific articles to be sure what is meant when the term "bar" is used.

94. Density. The density D of a substance is defined as the mass per unit of volume of that substance. If M is the total mass and V the total volume of the piece of the substance under consideration, then in symbols,

$$D \equiv \frac{M}{V} \tag{107}$$

and on clearing fractions,

$$M = DV. \tag{108}$$

Units of density will obviously be derived from the units in which M and V are expressed. Consequently,

When cgs units are used, density will be in $\dfrac{\text{grams}}{\text{cm}^3}$; and

When fps units are used, density will be in $\dfrac{\text{pounds}}{\text{ft}^3}$.

The **density of water** is used more frequently than that of any other substance. Its value comes directly from the definition of a gram of mass (Sec. 8); for 1 gm of mass equals the mass of 1 cm³ of water at 4°C, with sufficient accuracy for practical purposes. If we put these values of the mass and volume of this piece of water in Eq. (107), we have:

In cgs units the density of water is

$$D_w \doteq \frac{1 \text{ gm mass}}{1 \text{ cm}^3} \doteq 1 \text{ gm/cm}^3.$$

From this, by an easy computation or by experiment, we find that in fps units the density of water is $D_w = \dfrac{62.4 \text{ lb mass}}{1 \text{ ft}^3}$ $= 62.4 \text{ lb/ft}^3$. (For most purposes it is sufficiently accurate to take $D_w = 62.5 \text{ lb/ft}^3$.)

The densest element is osmium, having a density of 22.5 gm/cm³. The densest body is the "dark companion" of the star, Sirius. Its density is 1700 pounds per cubic inch.

95. Hydrostatic pressure is always normal to the restraining surface. When a fluid is at rest, its pressure against the containing vessel, or against any body in the fluid, is normal to the restraining surface at every point.

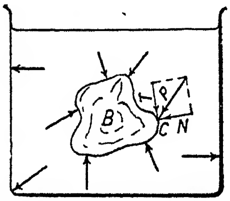


Fig. 86. Hydrostatic Pressure

This is easily proved. Let C be any point on the submerged surface of a body B in a fluid at rest (Fig. 86). Suppose that upon a very small, thin layer of the liquid surrounding C the resultant force P is not normal to the surface. Then we could resolve P into a normal component N and a tangential component T. The component N is balanced by the reaction of the body, but T is unbalanced; and since a fluid will yield to any unbalanced force, T will cause a flow parallel to the surface. But this is contrary to the hypothesis that the fluid is at rest.

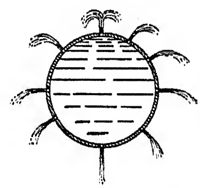


Fig. 87. Pressure Normal to Surface

Hence the pressure of a fluid at rest is always normal to the restraining surface. Such pressure is called **hydrostatic, or fluid, pressure.**

This theorem is verified by observation of water issuing from small holes drilled in a pressure line. The streams all leave the pipe normal to its surface, as shown in Fig. 87.

96. The free surface of a liquid at rest is horizontal. Consider a small volume of the liquid about a point C at the surface of a liquid at rest (Fig. 88). If the surface is not horizontal, the weight W of this little volume of liquid could be resolved into components N and T, normal and parallel respectively to the surface. Since the little volume is at the free surface, the pressure of the external air is normal to that surface (by Sec. 95) and merely increases N, which is counterbalanced by the reaction of the neighboring liquid. But T remains unbalanced,

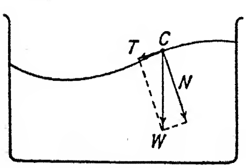


Fig. 88

and so will cause the little volume to flow downhill, which is contrary to the hypothesis that the liquid is at rest.

Hence the free surface must be horizontal.

In the foregoing, we have assumed gravity to be the only force acting. But it may be shown by similar reasoning that, in any case, a liquid surface which is free to do so will adjust itself so as to be **normal to the resultant pressure at every point.** This accounts for the curvature upward where water comes in contact with glass.

97. Pressure at any point within a fluid at rest is equal in all directions. Consider a tiny sphere of the fluid about any point C (Fig. 89), to be solidified without change of density. Suppose the force in some direction, say P_4, is greater than that in any other direction. Then, since a fluid yields to any unbalanced force, the little sphere would move in the direction of P_4, which is contrary to the hypothesis that the fluid is at rest. This proves either that the force, and hence the pressure, on the little sphere is equal in all directions; or else that the forces are balanced in pairs, thus: $P_1 = P_2$, $P_3 = P_4$, etc. Experiment confirms the former alternative, and hence the theorem is justified.

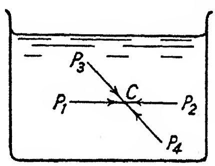


FIG. 89. Pressure within a Liquid

98. Pressure due to the weight of a liquid is proportional to the density of the liquid times the depth of the point below the surface of the liquid. Consider any point C at a depth h below the surface of the liquid whose density is D, the liquid being at rest.

Let us think of an area 1 cm^2, placed horizontal with its center at C. We may then think of the prism of water of height h cm, which stands on this 1-cm^2 base, as being divided by horizontal planes into h cubes, 1 cm on edge.

The mass of each cm^3 is by definition the density D of the liquid. Hence the mass of the prism (h cubes) is:

$$m = hD \text{ gm mass.}$$

Since, in gravitational systems, the weight of a unit of mass is a unit of force, the weight w of the prism of cubes sustained by the 1-cm^2 base at C is:

$$w = hD \text{ gm force.}$$

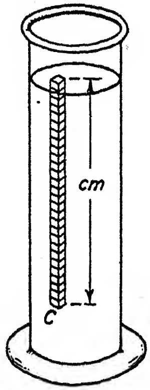


FIG. 90. Pressure within a Liquid

Since 1 gm of force = g dynes, we have also:

$$w = hDg \text{ dynes.}$$

But this force is acting upon unit area (1 cm)^2, and consequently is the pressure P, by Sec. 93. Therefore,

$$P = hD \text{ grams force/cm}^2 \tag{109}$$

$$= hDg \text{ dynes/cm}^2\text{, or baryes.} \tag{110}$$

If the above argument is carried out with British units, Eq. (109) obviously gives the pressure in pounds force/ft^2; and Eq. (110) gives it in poundals/ft^2.

Solved Problem

Find the pressure due to a column of water 1 ft high.
Since the density of water is 62.5 lb/ft^3, Eq. (109) becomes:

$$P = hD = 1 \text{ ft} \times 62.5 \text{ lb/ft}^3 = 62.5 \text{ lb/ft}^2$$
$$= \frac{62.5 \text{ lb/}\cancel{\text{ft}^2}}{144 \text{ in}^2/\cancel{\text{ft}^2}} = 0.434 \text{ lb/in.}^2$$

This fact, that the pressure due to water 1 ft deep is 0.434 lb/in.^{2}, will be found most useful in the solution of problems.

99. The foregoing law accounts for the well-known fact that **a liquid at rest stands at the same height (i.e., level) in all connected vessels, regardless of the size and shape of the vessels and of the connecting pipes.**

For example, let h_1, h_2, h_3 be the heights of a liquid in any vessels 1, 2, 3, respectively, above any given horizontal line MN (Fig. 91).

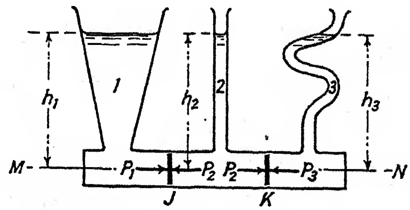


FIG. 91. Liquid in Connecting Vessels

Let frictionless disks J and K divide the apparatus into three compartments as shown.

Then, by Eq. (109), in gravitational units $P_1 = h_1D$; $P_2 = h_2D$; $P_3 = h_3D$. If A is the area of disk J, the force to the right on J is $= h_1DA$, and the force to the left on J is $= h_2DA$; and these two forces are equal since the liquid, and therefore the disk, is at rest.

Hence,

$$h_1DA = h_2DA$$

and

$$h_1 = h_2.$$

Similarly,

$$h_2 = h_3.$$

That is, the heights in all the vessels are the same, which was to be proved.

In slightly different form, this phenomenon is known as "the hydrostatic paradox."

100. Pascal's principle. This fundamental law of hydrostatics, given by Blaise Pascal in 1653, is as follows:

Pressure applied to an enclosed fluid is transmitted equally and undiminished in all directions to every part of the fluid and of its restraining surfaces.

Although true for all fluids, the law is easily proved in the case of an ideal, weightless, incompressible liquid.

Consider in Fig. 92 an ideal machine consisting of two pistons fitted to move without friction in two connected cylinders of areas A_1 and A_2, respectively, which are filled with the ideal, weightless, incompressible liquid.

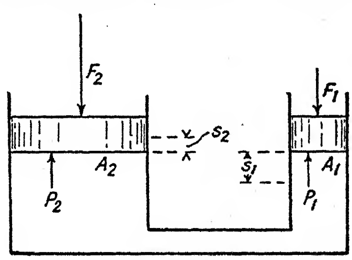


FIG. 92. Pascal's Principle

If a force F_1 depresses A_1 a distance s_1, a volume A_1s_1 of the liquid is forced over into the large cylinder, whose piston must therefore rise a distance s_2, such that

$$A_2s_2 = A_1s_1 \tag{a}$$

since the liquid is incompressible.

Again, the work done on the machine by F_1 must equal the work done by the machine against the load F_2, since it is assumed

that no energy is stored in the machine. Hence, by Eq. (82),

$$F_2 s_2 \times 1 = F_1 s_1 \times 1. \tag{b}$$

Dividing Eq. (b) by Eq. (a),

$$\frac{F_2 \cancel{s_2}}{A_2 \cancel{s_2}} = \frac{F_1 \cancel{s_1}}{A_1 \cancel{s_1}}.$$

But

$$\frac{F_2}{A_2} \equiv P_2 \quad \text{and} \quad \frac{F_1}{A_1} \equiv P_1.$$

Therefore,

$$P_2 = P_1.$$

But A_2 is representative of any surface in or bounding the liquid; hence the pressure P_1, produced by piston A_1, is transmitted undiminished as stated in the theorem.

Many applications of Pascal's principle will at once suggest themselves: the tire pump, the hydraulic jack, transmission of steam pressure from boiler to engine, etc.

101. Total force on a submerged surface. The total force on a submerged surface consists of two parts:

1. *Force F_1 due to liquid alone.* It is easily shown† that when a surface of area A, having any shape, is submerged in a liquid of density D, the force on the surface due to the liquid is:

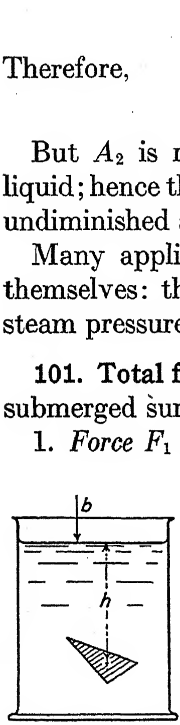


FIG. 93. Force on Submerged Surface

$$\left.\begin{aligned} F_1 &= AhD \quad \text{gravitational units} \\ &= AhDg \quad \text{absolute units} \end{aligned}\right\} \tag{111}$$

where h is **the depth of the center of gravity (centroid)** of the submerged surface below the surface of the liquid.

2. *Force F_2 due to atmospheric pressure.* The pressure b due to the atmosphere upon the surface of the fluid is transmitted to the submerged surface in accordance with Pascal's principle. Hence the force on the submerged surface due to atmosphere is:

$$F_2 = Ab. \tag{112}$$

† Mansfield Merriman, *Treatise on Hydraulics* (New York, John Wiley & Sons, 1916), Sec. 10.

The total force F on the submerged surface is, therefore,

$$F = AhD + Ab \text{ gravitational units.} \tag{113}$$

Example. What is the total force on the upstream face of the dam shown in cross section in Fig. 94, if the dam is 40 ft long (= width of river)?

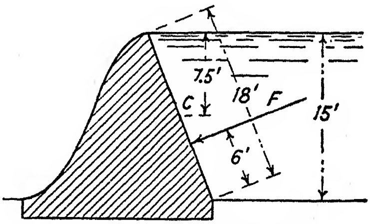


FIG. 94. Force on Dam

Since the atmospheric pressure on the water, and transmitted to the upstream side of the dam, is counterbalanced horizontally by the atmospheric pressure on the downstream side of the dam, we are here concerned with the force due to the water only.

The total force due to the water is by Eq. (111),

$$F = AhD.$$

where

$$A = 18 \times 40 = 720 \text{ ft}^2$$
$$h = \tfrac{1}{2} \times 15 = 7.5 \text{ ft (since the center of gravity of the submerged face of the dam is halfway down the face)}$$
$$D = 62.5 \text{ lb/ft}^3.$$

Hence

$$F = 720 \times 7.5 \times 62.5 = 337{,}500 \text{ lb force.}$$

This resultant force F acts perpendicularly to the face of the dam, by Sec. 95, but not at the center of gravity C of the upstream face, as might be expected. Its point of application is one-third of the way up the wetted face if that is rectangular.*

102. Archimedes' principle: A body is buoyed up by a force equal to the weight of the fluid which it displaces. The buoyant force acts vertically upward through the center of gravity of the displaced fluid.

For simplicity, consider a cube (Fig. 95) submerged in a liquid of density D, with two faces of the cube horizontal. Then the forces on the vertical faces are balanced in pairs, since opposite faces are at the same depth below the surface.

Then, by Eq. (111), we have:

Downward force on top face of cube $= l^2h_1Dg$ in absolute units.
Upward force on bottom face of cube $= l^2h_2Dg$ in absolute units.

* See Mansfield Merriman, *Treatise on Hydraulics* (New York, John Wiley & Sons, 1916), Secs. 12 and 13.

The buoyant force is the excess of the force upward over the force downward, therefore

$$\begin{aligned}\text{Buoyant force} &= l^2h_2Dg - l^2h_1Dg\\ &= l^2D(h_2 - h_1)g\\ &= l^2Dlg = l^3Dg.\end{aligned}$$

But $l^3 \equiv V$, the volume of the cube; and $VD = M$, the mass, by Eq. (108). Hence,

$$\text{Buoyant force} = VDg = Mg \text{ in absolute units.}$$

But Mg is the weight of the cube of liquid displaced by the body, by Eq. (56). So the body is buoyed up by a force equal to the weight of the liquid displaced.

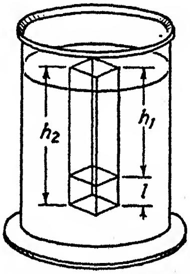


FIG. 95. Archimedes' Law

While this proof is given for a liquid, the law is equally true of any fluid, though the proof would be somewhat longer.

It is easily seen that the resultant buoyant force of the fluid acts vertically upward through the center of gravity of the displaced fluid. For, suppose the cube in Fig. 95 to consist of the fluid itself. Then the weight of the fluid displaced is the same as the weight of the cube of fluid; and the two just balance each other, since by hypothesis the cube of fluid is at rest.

But the weight of the cube is vertically downward through its center of gravity. Hence, **the buoyant force must be vertically upward through the same center of gravity.** Otherwise there would be a couple acting on the cube, which would produce turning, and this is contrary to the hypothesis that the fluid is at rest.

The center of gravity of the displaced fluid is called the *center of buoyancy*.

103. Determination of volume. It is often difficult if not impossible to determine the volume of a body of irregular shape, such as an apple, with the usual measuring instruments. In such cases the following methods may be used. They depend upon the fact that **when a body is submerged in a fluid, the volume of the fluid displaced is the same as the volume of the body.**

(*a*) *Volume by displacement.* Pour into a graduated cylinder an amount of liquid sufficient to submerge the body, and note the

volume. Submerge the body in the liquid and note the new volume. The difference is obviously the volume of the body.

If a sufficiently large graduate is not available, fill any vessel level full of a liquid, and then submerge the body in it. The amount that overflows is the volume of the body.

(*b*) *Volume by weighing.* In consequence of Archimedes' principle we have the following very accurate method. Let

$$\begin{aligned} W_a &= \text{weight of body in air, and} \\ W_l &= \text{weight of body in a liquid} \\ D_l &= \text{density of the liquid.} \end{aligned}$$

Then $W_a - W_l$ = apparent loss of weight in the liquid
= weight of liquid displaced (by Archimedes' principle)
= mass of liquid displaced if gravitational units are used.

By definition,

$$\text{Density } D \equiv \frac{M}{V}.$$

Therefore,

$$V = \frac{M}{D} = \frac{W_a - W_l}{D_l}. \tag{114}$$

Hence, **the volume of a body is its apparent loss of weight in a liquid divided by the density of the liquid.** Since the density of water is approximately 1 gm/cm^3, the volume of a body in cubic centimeters is practically equal to its apparent loss of weight in water, in grams.

104. Specific gravity. The density of a substance has been defined as its mass per unit of volume.

$$D \equiv \frac{M}{V}.$$

Two cubic feet of cast iron weigh about 900 lb. Hence, in the fps system,

$$\text{Density of cast iron } = \frac{900 \text{ lb}}{2 \text{ ft}^3} = 450 \text{ lb/ft}^3.$$

In the cgs system,

$$\text{Density of cast iron} = \frac{900 \times 453.6 \text{ gm}}{2 \times (30.48)^3 \text{ cm}^3} = 7.2 \text{ gm/cm}^3.$$

From this it will be seen that **density is always a compound number** and is not the same in the fps and cgs systems of units.

In order to have a quantity which is the same in both systems and which conveys essentially the same information as does density, we define specific gravity:

The specific gravity of a substance is the ratio of the density of the substance D_s to the density of water D_w at 3.98°C.

$$\text{Specific gravity (sp. gr.)} \equiv \frac{D_s}{D_w}. \tag{115}$$

From Sec. 94 we know that

$$\text{Density of water, } D_w = 62.4 \text{ lb/ft}^3 = 1 \text{ gm/cm}^3.$$

Hence, considering cast iron again, we find that,

In the fps system,

$$\text{Specific gravity of cast iron} = \frac{450 \text{ lb/ft}^3}{62.4 \text{ lb/ft}^3} = 7.2 \text{ (a pure number)}.$$

In the cgs system,

$$\text{Specific gravity of cast iron} = \frac{7.2 \text{ gm/cm}^3}{1 \text{ gm/cm}^3} = 7.2 \text{ (a pure number)}.$$

Thus it is seen that the specific gravity of a substance is always a **pure number** and has the same value in both systems of units.

In the cgs system, whenever we divide the density of a substance by the density of water, the denominator will always be 1 gm/cm³. Therefore, in the cgs system, specific gravity is numerically equal to density.

It should be carefully noted, however, that this is a **numerical equality only,** not a physical equality, for density is always a compound number while specific gravity is always a pure number; i.e., their dimensional formulas are not the same.

Density from specific gravity. It is frequently necessary to determine the density of a substance whose specific gravity is

known. For this purpose, we clear fractions in Eq. (115) and obtain the relation:

$$D_s = \text{Sp. gr.} \times D_w. \tag{115a}$$

TABLE OF SPECIFIC GRAVITIES

Air	0.001293 (0°C)	Balsa Wood	0.12
Alcohol, ethyl	0.79 (0°C)	Cork	0.25
Benzene	0.90 (0°C)	Ebony	1.22
Mercury	13.596 (0°C)	Oak	0.75
Water, sea	1.025 (15°C)	Pine	0.50
Brick	1.8	Aluminum	2.7
Cement	2.8	Copper	8.9
Glass	2.6	Lead	11.3
Ice	0.92	Osmium	22.5
Marble	2.7	Steel	7.6
Wax, sealing	1.8	Silver	10.5

105. Methods of determining specific gravity.

(*a*) *By weighing in air and in a liquid.*

Let W_a = weight of body in air
W_l = weight of body in liquid
V = volume of body
S_l = specific gravity of liquid
D_w = density of water.

Then $W_a - W_l$ = apparent loss of weight in the liquid
= mass of liquid displaced (if gravitational units are used)
= mass of liquid having volume V.

Hence

$$\text{Density of body} = \frac{W_a}{V}$$

and

$$\text{Density of liquid} = \frac{W_a - W_l}{V}$$

so that

$$S_l = \frac{\dfrac{W_a - W_l}{V}}{D_w}; \quad \text{from which } V = \frac{W_a - W_l}{D_w S_l}.$$

$$\text{Sp. gr. of body} \equiv S_b = \frac{\frac{W_a}{V}}{D_w} = \frac{W_a}{D_w\left(\frac{W_a - W_l}{D_w S_l}\right)}$$

$$= \frac{W_a}{W_a - W_l} S_l. \qquad (116)$$

FIG. 96. Specific Gravity Bottle

Thus, the specific gravity of a body may be found by dividing its weight in air by its loss of weight in *any* liquid and multiplying the quotient by the specific gravity of the liquid used. Since the specific gravity of water is approximately unity, **the specific gravity of a body is approximately equal to its weight in air divided by its apparent loss of weight in water.**

If the body does not sink in the liquid by its own weight, a sinker sufficient to submerge it must be attached. If W_s is the weight of the sinker alone in the liquid then

$$W_a + W_s = \text{weight of body in air and sinker in liquid;}$$

and if

$$W_{bs} = \text{weight of both body and sinker in liquid,}$$

$$(W_a + W_s) - W_{bs} = \text{loss of weight of body in liquid.}$$

(*b*) *By a specific gravity bottle, or pycnometer* (Fig. 96).

Let volume of bottle at 20°C = V ml
Mass of bottle empty = 12.362 gm
Mass of bottle filled with water at 20°C = 37.318 gm
Mass of bottle filled with H_2SO_4 at 20°C = 57.362 gm.

Then

Mass of water to fill bottle = 37.318 − 12.362 = 24.956 gm.

Similarly,

Mass of H_2SO_4 to fill bottle = 57.362 − 12.362 = 45.000 gm.

Therefore,

$$\text{Density of } H_2SO_4 \equiv D_b \equiv \frac{M}{V} = \frac{45.000}{V} \text{ gm/ml}$$

$$\text{Density of water} \equiv D_w = \frac{24.956}{V} \text{ gm/ml}$$

and

$$\text{Sp. gr. of } H_2SO_4 \equiv \frac{D_b}{D_w} = \frac{\frac{45.000}{V}}{\frac{24.956}{V}} = 1.8032 \text{ at } 20°\text{C in terms of water at } 20°\text{C.}$$

To correct for the variation of the density of water with temperature, it will be noted from tables that at 20°C water is a liquid whose specific gravity is 0.99823. Hence, by Eq. 116, the corrected value is

$$\text{Sp. gr. of } H_2SO_4 \text{ at } 20°\text{C} = 1.8032 \times 0.99832 = 1.800, \text{ referred to water at } 3.98°\text{C.}$$

(*c*) *By a density bulb, or sinker.* This is a ready method of determining specific gravity with apparatus that is always available. All that is required is a heavy bulb, or sinker, and a means of suspending it in the liquid. Sinker and suspension must not be acted upon chemically by the liquid. The Mohr, or Westphal, balance is designed especially for rapid determinations by this method, which approaches the accuracy of the specific gravity bottle method.

Let

V = volume of sinker
W_a = weight of sinker in air
W_w = weight of sinker in water
W_l = weight of sinker in other liquid.

Then

$W_a - W_w$ = apparent loss of weight in water
= mass of water displaced, if gravitational units are used,

and

$W_a - W_l$ = apparent loss of weight in other liquid
= mass of other liquid displaced, if gravitational units are used.

Hence,

$$\text{Density of other liquid} \equiv D_l = \frac{W_a - W_l}{V}$$

and

$$\text{Density of water} \equiv D_w = \frac{W_a - W_w}{V}$$

so that

$$\text{Sp. gr. of liquid} \quad = \frac{D_l}{D_w} = \frac{W_a - W_l}{W_a - W_w}. \qquad (117)$$

That is, if a sinker is weighed in air, in water, and in a second liquid, the specific gravity of the second liquid is equal to the ratio of the loss of weight in the second liquid to the loss of weight in water.

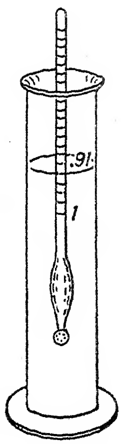


Fig. 97. Hydrometer

(*d*) *By the constant mass hydrometer.* Rapid and fairly accurate determinations of specific gravity may be made by means of the hydrometer (Fig. 97). This consists of a glass bulb having a scale tube at the top and weighted at the bottom so that it floats with its axis vertical. Since its mass is constant, the specific gravity of any liquid is the ratio of the volume displaced by the hydrometer when floating in water to the volume displaced when floating in the liquid to be tested. It is calibrated so that the specific gravity of the liquid in which it is floating may be read directly from the scale at the surface of that liquid.

106. Atmospheric pressure. The ocean of air at the bottom of which we live forms a gaseous envelope about the earth and is called the atmosphere. The molecules of the gases which compose the atmosphere are attracted by the earth, like all other particles, and hence **air has weight.** This may be verified by weighing an automobile tire before and after it has been pumped up.

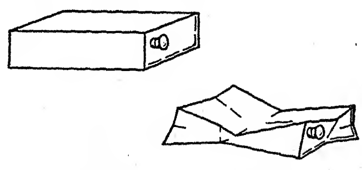


Fig. 98. Can Crushed by Atmospheric Pressure

Since air has weight, it must exert hydrostatic pressure upon all bodies in it, just as water or any other fluid does. This is clearly demonstrated by pouring a small amount of water into a gallon can and boiling it vigorously for a few minutes until the steam has replaced most of the air that the can originally contained. If the opening is then closed tightly with a rubber stopper and the can is allowed to cool, the steam condenses, causing a reduction of pressure inside until the greater pressure outside is sufficient to crush the can, as in Fig. 98.

Atmospheric pressure is utilized for many familiar purposes. When one sucks a liquid through a straw, or an animal drinks "up hill," he really reduces the pressure in his mouth by expanding the lungs. The liquid is then forced up by atmospheric pressure. Fountain pens and hydrometers are filled similarly, and pumps employ the same principle. The automatic chicken fountain and the oil supply of oil stoves are prevented from overflowing by atmospheric pressure on the outer surface of the liquid.

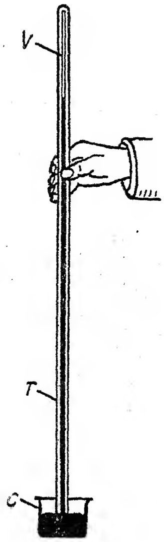


FIG. 99. Torricelli's Experiment.

107. Torricelli's experiment. Aristotle taught that nature "abhorred a vacuum" and filled immediately with the most available material any space in which a partial vacuum had been formed. This belief prevailed until Galileo was called as consulting engineer to determine why a suction pump would not raise the water from a deep well belonging to his friend, the Duke of Tuscany. Observing that the water was raised only about 28 feet, Galileo remarked that apparently there was a limit to nature's abhorrence of a vacuum.

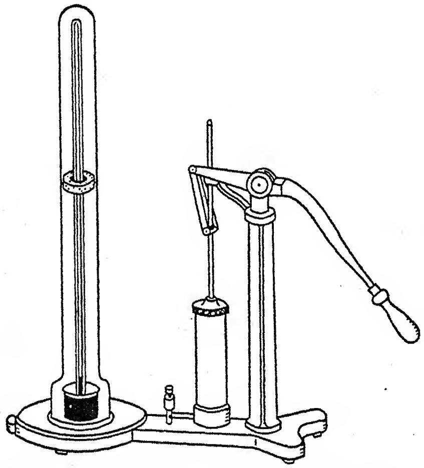
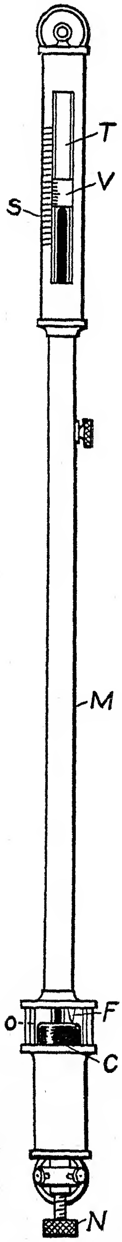


FIG. 100. Barometer in Vacuum

After Galileo died in 1642, Torricelli, one of his pupils, continued the investigation. He filled with mercury a glass tube closed at one end. With a finger over the open end, he inverted the tube and placed its open end beneath the surface of mercury in a bowl, without allowing air to enter the tube. When his finger was removed, the mercury descended in the tube until the column sustained was about 30 in. high (Fig. 99). Torricelli felt satisfied that the column was supported by atmospheric pressure; but the con-

clusive evidence was obtained by Pascal, who had one of the tubes taken to the top of Puy de Dome, a mountain in southern France. There, at a height of 1000 meters, the column was nearly 7 cm shorter than at Paris.



Today, any doubt that the column of mercury is supported by the pressure of atmosphere may be dispelled by placing the tube and basin of mercury in a tall bell jar (Fig. 100) which is connected to an air pump. As the air is exhausted, the mercury descends. When air is re-admitted to the large tube, the mercury returns to its original height in the small tube.

108. The Torricellian vacuum. Above the mercury in the tube of Fig. 99, the space V is largely devoid of matter. This is called a "Torricellian vacuum." Though a very good vacuum, it is far from perfect, because molecules of mercury evaporate into this space until they produce the vapor pressure of mercury corresponding to the temperature of the apparatus. Much better vacuums may be produced by modern vacuum pumps.

109. The mercurial barometer (**Fortin's**). In improved form, Torricelli's apparatus is today our **most reliable method of measuring atmospheric pressure.** The glass tube T (Fig. 101) and cistern of mercury C are encased in a metal tube M on which is engraved the scale S. The zero point of this scale is at the tip of a small glass finger, or index, F, which projects downward above the mercury in C.

Before taking a reading, the surface of the mercury must be adjusted by means of the screw N until it is just tangent to the index, F. The height of the column is then read on the scale with the aid of the movable vernier V.

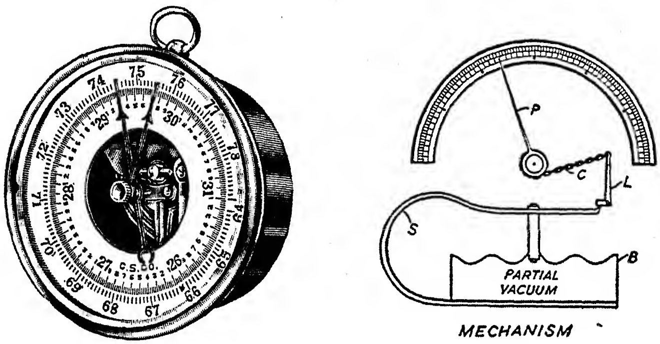
110. The aneroid barometer. This instrument, while less accurate than the mercurial barometer, has

FIG. 101. Mercurial Barometer. (Courtesy Welch Scientific Co.)

the great advantage of ruggedness and portability, and its accuracy is amply sufficient for much practical work.

As is shown in Fig. 102, the mechanism consists of a cylindrical box B having a corrugated cover and partially exhausted of air. To this is fastened the spring S, which, by means of the train of levers and the chain C, actuates the pointer P.

When atmospheric pressure increases, the top of B is depressed and the pointer moves clockwise, and vice versa. The instrument is subject to errors due to temperature, changes of elasticity of the spring, and friction of the mechanism. It must be corrected by comparing simultaneous readings taken on it and on a mercurial barometer.



(a) (b)

FIG. 102. Aneroid Barometer (Courtesy Central Scientific Co.)

111. Pressure of one standard atmosphere. A pressure of one standard atmosphere is defined as the pressure exerted by a column of mercury 76 cm high at 0°C, 45° latitude, and mean sea level.

This is not, of course, the actual pressure of the atmosphere at any particular place and time. The pressure of the atmosphere at different places depends upon the latitude and the altitude of the place, and also varies from hour to hour with the weather conditions. As a storm approaches the pressure as indicated by the barometer falls, whereas a rising barometer indicates the approach of fair weather.*

* See Government Weather Map.

But averaged over a long period of time, the pressure at sea level and at 45° latitude is about equal to that produced by a column of mercury 76 cm high at 0°C. This value is accordingly adopted as a standard.

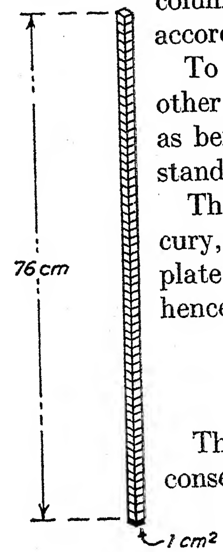


FIG. 103

To express the pressure of a standard atmosphere in other units, let us think of this column of mercury as being 1 square centimeter in cross section and as standing on a plate D of that area (Fig. 103).

Then, for each cm of height, there is 1 cm^3 of mercury, or 76 cm^3 in the whole column standing on the plate. The density of mercury at 0°C is 13.6 gm/cm^3, hence the column of mercury weighs

$$76\,\cancel{cm^3} \times 13.6\,\frac{\text{gm force}}{\cancel{cm^3}} = 1033.6 \text{ gm force.}$$

This weight of mercury stands on an area of 1 cm^2, consequently,

(a) The pressure of 1 standard atmosphere $= 1033.6\ gm/cm.^2$

Since 1 gm is (very nearly) the weight of 1 cm^3 of water, to produce a pressure of 1033.6 gm/cm^2 would require a column of 1033.6 cm^3 of water to stand on our 1-cm^2 plate. This column of water would be 1033.6 cm high, hence

(b) The pressure of 1 standard atmosphere $=$ 1034 cm of water
$=$ 33.9 ft of water.

From (Sec. 98), a column of water 1 ft high produces a pressure of 0.434 $lb/in.^2$ Therefore,

(c) The pressure of 1 standard atmosphere $= 33.9 \times 0.434$
$= 14.7\ lb/in.^2$

In the preceding, we have deduced, from the definition of a standard atmosphere, its equivalent in the various units in which it is most frequently expressed. The definition and the last value should be memorized.

The pressure of the atmosphere at any point is due to the weight of the column of air, say, 1 cm^2 in cross section, extending from that point up to the top of the atmosphere (approximately 100 miles). Consequently, its pressure at the top of a table will be

less than that at the floor; and a good aneroid barometer will indicate this difference.

112. Standard conditions. When the results of different experimenters are to be compared, it is obviously necessary that they be reduced to what they would have been had all the experimenters worked under the same conditions of temperature, pressure, etc.

Standard conditions are defined as 0°C, 76 cm of mercury pressure, and $g_n = 980.665$ cm/sec^2.

HYDRAULICS

113. Hydraulics treats of fluids in motion. A fluid which flows so that its velocity at each point is constant in magnitude and direction is said to have simple, or **steady flow.** Thus if, as the flow continues, a particle which at any instant has reached point A always has the velocity v_1 and a particle which has reached point B always has the velocity v_2, etc., the liquid is said to have simple flow. The path ABC followed by a particle of the fluid in steady flow is called a stream line. Stream lines (light lines in Fig. 104) obviously never cross one another, for the particles at the intersections would then have to move in the direction of both stream lines at once, which is impossible.

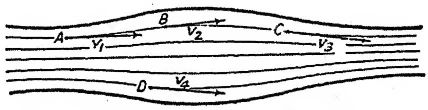


FIG. 104. Steady Flow

114. Tube of flow. We may select a sequence of continuous stream lines which form the elements of a tubular surface within

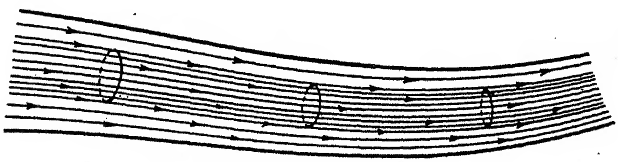


FIG. 105. Tube of Flow

the fluid. That part of the fluid bounded laterally by such a surface is called a *tube of flow* (Fig. 105).

Since stream lines do not cross and a particle continues along a single stream line, every particle within a given tube of flow will

continue within that tube. That is, there is no passing of particles through the lateral surface of a tube of flow.

115. With steady flow the volume V of a fluid that passes any plane cross section of a tube per sec is equal to the area A of the cross section times the average velocity v of particles at the cross section (Fig. 106).

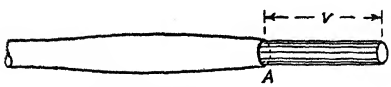


FIG. 106. Volume Discharged

For, imagine the fluid to congeal, without change of density, as it passes the cross section A. Then in 1 second there would emerge a right cylinder or prism of altitude v and area A. The volume of this cylinder would be, by geometry,

$$V = Av. \tag{118}$$

Since the flow is steady, this volume would pass any other cross section of the tube in 1 second.

Simple flow is an ideal state that can be realized only approximately. In actual cases, there are always eddies, or turbulence.

116. Law of continuity of flow of a liquid. When a liquid has steady flow, the velocity varies inversely as the cross section of the tube of flow.

In Fig. 107, let v_1, v_2, v_3, etc., be the mean velocities of the liquid where the cross-sectional areas are A_1, A_2, A_3, etc., respectively.

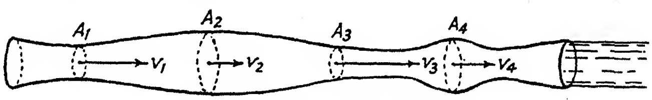


FIG. 107. Continuity of Flow

Since the liquid is assumed to be incompressible, the volume which passes any cross section A_1 per unit time must be the same as that which passes any other cross section A_2 in the same time; i.e.,

$$\text{volume per sec} = A_1v_1 = A_2v_2$$

from which, by algebra,

$$\frac{v_1}{v_2} = \frac{A_2}{A_1} \tag{119}$$

which was to be proved.

The theorem is true only for liquids which are practically incompressible.

In the case of **compressible fluids,** the density D is not constant but varies with the pressure. With steady flow, however, the mass M of the fluid that passes any cross section A_1 must equal that M_2 which passes any other cross section A_2 in the same time. Hence, for all fluids we may write:

$$M_1 = M_2 = M_3 = \ldots$$

But $$M = DV. \qquad \text{by Eq. (108)}$$

Hence, $$V_1D_1 = V_2D_2 = V_3D_3 = \ldots.$$

$$V = Av. \qquad \text{by Eq. (118)}$$

Therefore, $$A_1v_1D_1 = A_2v_2D_2 = A_3v_3D_3 = \ldots. \qquad (120)$$

for any fluid whatever.

117. Energy transmitted by a liquid due to pressure. Consider the mechanism of Fig. 108, horizontal and filled with a frictionless, incompressible fluid, and operating with infinite slowness.

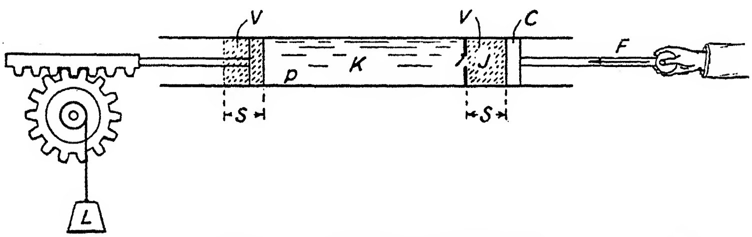


FIG. 108. Liquid Transmits Energy

Let the hand exert upon the piston C a gradually increasing force F. No liquid will flow into tank K until F becomes great enough to produce a pressure p dynes/cm² in pump J equal to that in K; i.e., until

$$F = pA \text{ dynes},$$

where A cm² is the area of the piston C. The valve will then open; and if C advances a distance s cm, the volume of liquid forced into K will be:

$$V = As \text{ cm}^3$$

and the total work W required to get this volume into the tank will be:

$$W = Fs \cos \phi \qquad \text{by Eq. (85)}$$

and since $\phi = 0$,

$$W = pAs1$$
$$= pV \text{ ergs.}$$

Thus, the **work done per unit volume** of the liquid pumped into K will be:

$$w \equiv \frac{W}{V} = \frac{pV}{V} = p \text{ ergs/cm}^3. \tag{121}$$

Or, if the pressure in K were p' gm force/cm², the work per unit volume would be:

$$w = p' \text{ gm-cm/cm}^3. \tag{122}$$

Consider now the volume V of liquid that has been forced into the tank where the pressure p was maintained:

It has acquired no gravitational potential energy, because it has not been elevated.

It has acquired no kinetic energy, because the change took place with infinite slowness.

It has acquired no potential energy due to pressure, because it was incompressible.

No energy has been lost, because the apparatus and liquid are frictionless.

Hence it cannot properly be said that any energy has been stored in or lost by the liquid. The liquid has served only to transfer an amount of energy pV from the pumping system to the lifting system, just as a solid rod might have done. The result may therefore be stated as follows:

Energy transmitted by a liquid per unit volume due to pressure*

$$= p \text{ ergs/cm}^3 \quad \text{if pressure is } p \, \frac{\text{dynes}}{\text{cm}^2} \tag{123}$$

$$= p' \text{ gm-cm/cm}^3 \quad \text{if pressure is } p' \, \frac{\text{gm force}}{\text{cm}^2}. \tag{124}$$

118. Energy of a fluid in motion. Let a mass of fluid M have a density D and a velocity v.

* This energy is often erroneously called "pressure potential energy of the liquid per unit volume."

Its kinetic energy $= \frac{1}{2}Mv^2$ ergs. by Eq. (86)

Its volume, $V = \dfrac{M}{D}$ cm³ from the definition of density.

Therefore,

$$\textbf{Kinetic energy per unit volume} = \frac{\frac{1}{2}Mv^2}{V} = \tfrac{1}{2}Dv^2\,\frac{\text{ergs}}{\text{cm}^3} \qquad (125)$$

$$= \tfrac{1}{2}\frac{D}{g}\,v^2\,\frac{\text{gm-cm}}{\text{cm}^3}. \qquad (126)$$

119. Energy of a fluid due to its elevation. Consider a mass M (Fig. 109) of a fluid of density D at a mean height h above some arbitrary datum plane. Then, because of its position,

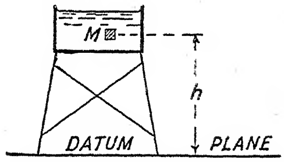


FIG. 109. Energy of Elevated Liquid

Its potential energy $= Mh$ gm-cm by Eq. (89)
$= Mgh$ ergs.

Its volume, $V = \dfrac{M}{D}$ cm³ from the definition of density.

Therefore,

$$\textbf{Potential energy per unit volume (due to its elevation)} = \frac{Mh}{\frac{M}{D}} = Dh\,\frac{\text{gm-cm}}{\text{cm}^3}$$

$$= Dgh\,\frac{\text{ergs}}{\text{cm}^3}. \qquad (127)$$

120. Bernoulli's principle. Let MN (Fig. 110) be a tube of flow in a frictionless, incompressible fluid having steady flow, and consider the definite section between end planes at any two points J and K on the same stream line.

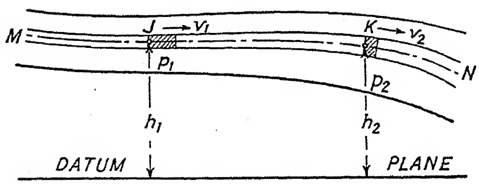


FIG. 110. Bernoulli's Principle

When 1 cm³ of fluid passes J it brings into the J–K section of the tube:

K.E. $= \frac{1}{2}Dv_1^2$ ergs/cm³ by Eq. (125)

Gravitational P.E. $= Dgh_1$ ergs/cm³. by Eq. (127)

It also transmits from the fluid behind it to the fluid in front of it, on account of the pressure,

$$\text{An amount of energy} = p_1 \quad \text{ergs/cm}^3. \qquad \text{by Eq. (123)}$$

Hence, for each cm³ of fluid entering at J, the total energy passing into the section J–K is:

$$E_1 = \tfrac{1}{2}Dv_1^2 + Dgh_1 + p_1 \quad \text{ergs/cm}^3. \tag{a}$$

According to Sec. 114, no fluid passes into or out from a tube of flow through its lateral surface. Hence, since the fluid is incompressible, when 1 cm³ is pushed in at J another cm³ must be pushed out at K.

The cm³ leaving the J–K section at K will carry out of the section:

$$\text{K.E.} = \tfrac{1}{2}Dv_2^2 \quad \text{ergs/cm}^3$$

$$\text{Gravitational P.E.} = Dgh_2 \quad \text{ergs/cm}^3$$

and it will transmit by pressure from the fluid behind it to the fluid in front of it

$$\text{An amount of energy} = p_2 \quad \text{ergs/cm}^3.$$

Therefore, the total energy passing out of the section per unit volume leaving at K is:

$$E_2 = \tfrac{1}{2}Dv_2^2 + Dgh_2 + p_2 \quad \text{ergs/cm}^3. \tag{b}$$

Since, by hypothesis, the flow is steady and the fluid is frictionless, the total energy of the section J–K must be constant. Hence, by the principle of conservation of energy, the energy received per unit volume entering at J must equal the energy passing out per unit volume leaving at K. That is

$$E_1 = E_2$$

or, from Eqs. (a) and (b), above,

$$\tfrac{1}{2}Dv_1^2 + Dgh_1 + p_1 = \tfrac{1}{2}Dv_2^2 + Dgh_2 + p_2.$$

But J and K were any two points on the same stream line. Hence the energy transmitted along a stream line per unit volume passing a given point on the line is:

$$\tfrac{1}{2}Dv^2 + Dgh + p = C, \text{ a constant (absolute units)} \tag{128}$$

$$\tfrac{1}{2}\frac{D}{g}v^2 + Dh + p' = C', \text{ a constant (gravitational units)} \tag{129}$$

where p is in dynes/cm^2 and p' is in grams force/cm^2; or both are in the corresponding British units.

This equation is the relation known as Bernoulli's principle. It was first given by Daniel Bernoulli (1700–1782) and is the most important theorem in hydromechanics.

Bernoulli's principle is often stated by engineers in a different form:

Dividing Eq. (129) through by D, we have:

$$h + \frac{p'}{D} + \frac{v^2}{2g} = \frac{C'}{D} \equiv H. \tag{130}$$

Since h is a height, or head, each of the other terms must also represent a height, by dimensional reasoning.* The engineer calls

h the *elevation head*

$\frac{p'}{D}$ the *pressure*, or *static*, *head*

$\frac{v^2}{2g}$ the *velocity*, or *kinetic*, *head*

H the *total head.*

He therefore says: *The total head along a stream line is constant.*

While rigorously correct only for incompressible, non-viscous liquids, Bernoulli's theorem may be applied to ordinary liquids with sufficient accuracy for most engineering purposes.

121. Applications of Bernoulli's theorem. Torricelli's theorem follows at once as a special case of Bernoulli's.

Consider a unit volume at the surface and at the orifice (Fig. 111a), and let the datum plane be taken through the center of the orifice.

Then at the surface,

$$\begin{aligned} h &= h \\ p' &= b = \text{atmospheric pressure} \\ v &= 0 \end{aligned}$$

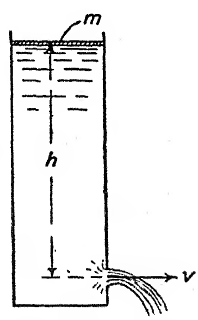


FIG. 111a. Torricelli's Theorem

* See Appendix I.

and at the orifice,

$$h = 0$$
$$p' = b = \text{atmospheric pressure}$$
$$v = v.$$

Putting these values into Eq. (130),

$$\begin{array}{cc} \text{At the surface} & \text{At the orifice} \end{array}$$

$$h + \frac{b}{D} + \frac{0}{2g} = 0 + \frac{b}{D} + \frac{v^2}{2g}$$

whence

$$\frac{v^2}{2g} = h$$

and

$$v = \sqrt{2gh} \tag{131}$$

which is Torricelli's theorem in algebraic form, and states that under the action of gravity alone the velocity of flow of a liquid from an orifice is the same as that of a particle falling freely from the surface of the liquid to the orifice. On account of contraction of the stream as it leaves the orifice, the actual volume discharged is only about 61% of the theoretical amount.

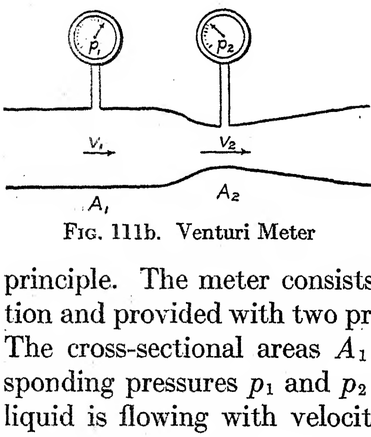


FIG. 111b. Venturi Meter

The *Venturi meter* for determining the rate of flow of a liquid in a pipe line affords a very practical application of Bernoulli's principle. The meter consists of a pipe having a constricted section and provided with two pressure gauges as shown in Fig. 111b. The cross-sectional areas A_1 and A_2 are known, and the corresponding pressures p_1 and p_2 are read from the gauges while the liquid is flowing with velocities v_1 and v_2, respectively, through these cross-sections.

It is required to find the volume V of the liquid that flows per second past any cross-section.

By Eq. 119 $$A_1v_1 = A_2v_2 \tag{a}$$

and by Bernoulli's principle,

$$0 + \frac{p_1}{D} + \frac{{v_1}^2}{2g} = 0 + \frac{p_2}{D} + \frac{{v_2}^2}{2g}. \tag{b}$$

Since in any given case, the density D and g will be known, Eqs. (a) and (b) may be solved for v_1 and v_2.

The volume flowing through the pipe per second is then gotten from the relation,

$$V = A_1 v_1.$$

PROBLEMS

1. In order to find the diameter of a capillary tube, it is weighed empty and when filled with mercury. If the mercury required to fill a tube 5 cm long at 20° weighs 0.8 gm, what is the diameter of the tube?

2. At 15°C the mercury required to fill a tube 8 cm long weighs 1.2 gm. What is the diameter of the tube?

3. What is the volume of a glass bulb that holds exactly 342 gm of mercury at 18°C?

4. In order to find the volume of the bulb of a McLeod gauge, it is found to weigh 52.36 gm empty and 1140.73 gm when filled with mercury at 20°C. What is the volume of the bulb at 20°C?

5. If the specific gravity of sea water is 1.025, what volume of water is displaced by a ship of 35,000 tons displacement?

6. A barge is 30 × 120 ft. How much deeper does it sink in salt water when loaded with 6 freight cars of 100,000 lb each?

7. A barge 16 ft wide and 40 ft long sinks 18 in. when a loaded car is run on it. What is the weight of the car and its contents?

8. What must be the volume of the bag of a balloon filled with hydrogen (density 0.00009 gm/cm^3) to lift a load of 500 kg, including the weight of the bag, if the density of the air is 0.001293 gm/cm^3?

9. If the densities of air, hydrogen, and helium are 0.001293, 0.000090, and 0.000177 gm/cm^3, respectively, how much would the lifting force of the Hindenburg have been reduced if it had been filled with helium instead of hydrogen, its volume being 7,063,000 m^3?

10. In the hydrostatic bellows, a horizontal bellows 15 in. square is connected to a vertical pipe. What must be the height of the water in the pipe above that in the bellows to balance a 150-lb man standing on the bellows?

11. An elliptical manhole in a standpipe has an area of 240 in.^{2} and its center is 80 ft below the surface of the water. With what force is it held on?

12. The upstream face of a dam is inclined 15° to the vertical. If the river is 140 ft wide and 12 ft deep, what is the total force on the face of the dam when filled?

13. If the upstream face of a dam is 300 ft long and 40 ft high, measured on the incline which is 20° to the vertical, what is the total force on the face of the dam when full of water?

14. A dam is built across a river 90 ft wide and 20 ft deep. If the upstream face of the dam is inclined 12° to the vertical, what is the total force upon it when filled with water?

15. A dam has a rectangular upstream face 200 ft long and 30 ft high meas-

ured on the incline, which is 10° to the vertical. What is the total force on the face of the dam when full?

16. A sinker weighs 16 gm in air and 10 gm in water. It weighs 8 gm in another liquid. What is the specific gravity of the liquid? Of the sinker?

17. The alloy used for silver coin is 90 parts silver and 10 parts copper by weight. What is the density of coin silver? *Ans.* 10.3 gm/cm^3.

18. A piece of cork weighs 2 gm in air. A sinker alone in water weighs 50 gm. When the sinker is attached to the cork and both are submerged in water, their combined weight is 44 gm. What are the volume and density of the cork?

19. When a piece of balsa wood is submerged in water by a sinker, together they weigh 24.57 gm. The wood alone in air weighs 0.57 gm, and the sinker alone in water weighs 30 gm. Find the specific gravity and volume of the wood.

20. A block of wood having a cross section 5 by 4 cm and a height of 3 cm floats in water immersed to a depth of 2.5 cm. What mass laid on top would be just sufficient to cause complete immersion of the block?

21. A piece of pine floats with 1/4 its volume out of water. What is its density?

22. If the specific gravity of ice is 0.92 and that of sea water is 1.025, how many cubic yards of an iceberg are below water when 1000 yd^3 are above water?

23. An iceberg in salt water floats with 0.9 of its mass submerged. What is the density of the ice if the specific gravity of the sea water is 1.025?

24. A specific gravity bottle weighs 26.4 gm empty. Filled with water at 20°C it weighs 76.2 gm, and filled with gasoline it weighs 63.9 gm. What is the density of the gasoline?

25. A piece of potassium weighs 0.21 gm in oil whose specific gravity is 0.8, and 0.51 gm in oil whose specific gravity is 0.7. Find the mass, specific gravity, and volume of the specimen.

26. A piece of sodium weighs 0.34 gm in oil whose specific gravity is 0.8, and 0.54 gm in oil whose specific gravity is 0.7. Find the mass, specific gravity, and volume of the specimen.

27. A piece of lithium weighs 0.068 gm in oil of specific gravity 0.5, and 0.268 gm in oil of specific gravity 0.4. Find the mass, volume, and density of the lithium.

28. A certain quantity of sugar weighs 80.5 gm in air and 45.5 gm in oil whose specific gravity is 0.7. What is the density of the sugar in pounds per cubic foot?

29. A certain quantity of salt weighs 43.6 gm in air and 28.4 gm in an oil whose specific gravity is 0.76. What is the density of the salt in pounds per cubic foot?

30. Batches of shortening are weighed in vats which stand on platform scales. Steampipes fastened to the ceiling are submerged in the shortening to keep it liquid. If the submerged volume of the pipes is 2 ft^3 and the specific gravity of the liquid is 0.85, what is the true weight of the batch when the apparent weight is 1000 lb?

31. A piece of soap weighs 18 gm in air and 2 gm in an oil whose specific gravity is 0.8. What is the density of the soap?

32. A geologist, to determine the height of a mountain, notes that at the foot

his aneroid reads 29.6 in. of mercury and at the top 25.6 in. If the mean density of the air is 0.078 lb/ft^3, what is the height of the mountain?

33. In a horizontal Venturi meter the pressure is 40 lb/in.^{2} where the diameter is 2 in., and 30 lb/in.^{2} where the diameter is 1 in. What is the rate of flow of the water in gallons per minute?

34. In a Venturi meter the pressure is 50 lb/in.^{2} where the diameter is 1 in., and 40 lb/in^2. where the diameter is 1/2 in. If the meter is horizontal, what is the rate of flow of water in gallons per minute?

35. The diameter of a horizontal jet pump is 3/4 in. where it connects to the water system, and 1/8 in. where the suction tube attaches. When the pressure in the mains is 40 lb/in.^{2} and the flow is 0.006 ft^3/sec, what is the pressure at the top of the suction tube?

36. One ft^3 of water per second are to be pumped in at the top of an open tank 100 ft high through a delivery pipe having an inside diameter of 2 in. If the mains have an inside diameter of 4 in., what must be the pressure in the mains, neglecting friction?

37. In the preceding problem, what must be the pressure in the mains if the water is pumped at the same height into a closed tank in which the pressure is 40 lb/in.^{2}?

38. At what rate in gal/min will water flow from a circular hole 4 in. in diameter and whose center is 60 ft below the water surface in a large tank, if the coefficient of discharge is 0.75?

39. If a circular hole 4 ft in diameter is blown in the side of a ship 15 ft below the water line, at what rate in gal/min will water enter the ship?

40. If the gauge pressure in a water line is 25 lb/in.^{2} when a spigot is open, at what rate in gal/min will water flow from a ½ in. spigot near the gauge? Assume the coefficient of discharge to be 0.75.

CHAPTER VIII

GRAVITATION

122. Universal gravitation. The story is told of Sir Isaac Newton (1642–1727) that, while lying under an apple tree, he was struck by a falling apple. Reflecting upon this familiar phenomenon, he recognized the really remarkable fact that apples always fall toward rather than away from the earth, and began to wonder if the earth likewise did not fall toward the apple. Whether or not his thoughts were focused on gravitation by so commonplace an experience, Newton succeeded in unifying the work of Copernicus, Galileo, and Kepler into one of the most brilliant generalizations in the history of science.

The law of universal gravitation states that: Every particle of matter in the universe attracts every other particle with a force F which varies directly as the product of their masses m_1 and m_2 and inversely as the square of the distance r between them. Stating this law in mathematical symbols,

$$F \propto \frac{m_1 m_2}{r^2}$$

from which, by the definitions in algebra,

$$F = G\frac{m_1 m_2}{r^2}. \tag{132}$$

The constant of proportionality G is called the *gravitational constant*, or the *Newtonian constant.* It is not a pure number and its value depends upon the choice of units for the other quantities. A value of G was first determined in 1778 by the Rev. Nevil Maskelyne, who observed the deviation from the vertical of a plumb line suspended near Mt. Schiehallion in Scotland.*

Spheres, solid or hollow, attract each other as if the entire mass of each were concentrated at its center of mass.† This is not true

* For this and other methods, see Poynting, "Gravitation," in *Encyclopedia Britannica*, 11th ed.

† It is said to have taken Newton twenty years to prove this fact.

for bodies having any form other than spherical (see Sec. 71). However, if the bodies are small compared to the distance between them, it is usually a satisfactory approximation.

Though not equipped with apparatus sufficiently sensitive to verify the law of gravitation in the laboratory, as we are now able to do, Newton satisfied himself of the universality of the law by showing that the gravitational force of the earth upon the moon was equal to the centripetal force necessary to hold the moon in its nearly circular orbit around the earth. Astronomers found that it accounted satisfactorily for the motions of the planets about the sun (except in the case of Mercury); and by its means Leverrier and Adams predicted and located the planet Neptune where it was subsequently discovered.

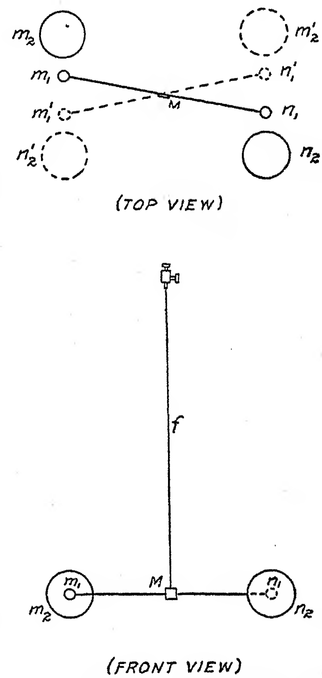


FIG. 112. Cavendish's Experiment

123. Cavendish's experiment. Henry Cavendish, first director of the physics laboratory at the University of Cambridge, England, in 1798 made the first successful laboratory determination of G. His method, with later improvements, is substantially as follows.

Two equal spheres of platinum, m_1 and n_1 (masses of about 1 gm each), are mounted at the ends of a light rod (about 2.5 cm long) which carries at its center a small mirror M (Fig. 112). This rod is suspended by a fine quartz fiber f from a bracket so as to hang symmetrically between two equal lead spheres m_2 and n_2 (masses of about 3 kg each).

When the lead balls are shifted from positions m_2, n_2 to dotted positions m_2', n_2', the gravitational attractions of the lead balls for the platinum balls change sides, so that the torque is changed from clockwise to counterclockwise and the suspended system is turned through a small angle into the dotted position. The angle may be measured with great precision by observing the displace-

ment along a scale of a spot of light reflected from the mirror M.

From the value of this angle, the torsional constant of the fiber, the masses and the dimensions of the apparatus, the gravitational attraction of a lead ball for a platinum ball can be computed, and the constant G found.

The latest determination of G was made in 1930 at the U.S. Bureau of Standards by Dr. Paul R. Heyl, who finds

$$G = 6.670 \times 10^{-8}\ \text{cm}^3\ \text{gm}^{-1}\ \text{sec}^{-2}$$

when cgs units are used.

124. Einstein's law of universal gravitation. Time and distance are such familiar concepts that Galileo and Newton based their mechanics implicitly upon the popular idea that these and mass are quantities absolute and independent in nature. But in 1905, at the University of Berlin, Albert Einstein, in endeavoring to explain the negative result of the Michelson-Morley experiment (Sec. 626), was led to question and discard that time-honored assumption.

Showing that length, mass, and time are related and that their values depend upon the circumstances under which they are measured, Einstein succeeded in giving a rational explanation of the Michelson-Morley experiment and evolved his *General Theory of Relativity*, the publication of which, in 1915, marked the beginning of a new epoch in scientific thinking.

Einstein's law of gravitation, one of the consequences of his theory of relativity, is more general than Newton's and accounts for the previously anomalous motion of the planet Mercury. But for its elucidation it requires mathematics quite beyond the scope of the usual college course; and its final expression, in six differential equations, is much less simple than that of Newton's. Consequently, Newton's law, which is a special case of Einstein's and is sufficiently accurate except in the case of a body moving with high velocity in a strong gravitational field, will continue to be generally used.

125. Variation of g. Consider a body of mass m_1 on or above the surface of the earth (mass m_2), and let r be its distance from the center of the earth. Then neglecting the effect of the earth's rotation, the weight of the body is approximately the gravita-

tional attraction of the earth for it. That is

$$\text{weight} \doteq F = G\frac{m_1 m_2}{r^2}.$$

But the weight of a body gives to it the acceleration g. Consequently by Newton's second law of motion

$$F \doteq m_1 g.$$

Hence
$$m_1 g \doteq G\frac{m_1 m_2}{r^2}.$$

Whence
$$g \doteq G\frac{m_2}{r^2}.$$

This value of g is what it would be at the distance r from the center of a perfectly homogeneous spherical earth which was not rotating. In that case it is seen that outside the earth the *acceleration of gravitation* of the earth varies inversely as the square of the distance from the center of mass of the earth to the center of mass of the other body, if that body also is a homogeneous sphere; otherwise this distance would be approximately correct.

For this reason, g at the equator is less than it is at the poles, since the earth's equatorial radius is greater than its polar radius; and the value of g at any place depends upon its elevation above mean sea level.

On account of the earth's rotation, bodies at all places except the poles tend to fly off along a tangent. This gives the effect of a small acceleration (3.36 cm/sec² at the equator) opposed to gravitation, which obviously depends upon the latitude of the place in question.

Helmert has given the following formula for g, which gives values that check closely with experiment:

$$g = 980.616 - 2.5928 \cos 2\lambda + 0.0069 \cos^2 2\lambda - 0.0003086\, H \qquad (133)$$

where λ is the latitude and H is the altitude in meters.

However, the density of the earth is not uniform, and where its denser constituents such as iron are near the surface, the value of g is greater than is given by Eq. (133). Where the greatest precision is required, g is determined by pendulum experiments.

VALUES OF ACCELERATION OF GRAVITY AT VARIOUS PLACES

Place	Value of g in cm/sec²
Standard, or Normal, g_n	980.665
At Equator (sea level), g_0	978.039
At 45° (" "), g_{45}	980.616
At Pole (sea level), g_{90}	983.217
At Washington, D.C. (Bureau of Standards)[a]	980.080
At Washington, D.C. (C. and G. S.)	980.092
At Potsdam, Germany (old basic value)	981.274

[a] Heyl and Cook, Bureau of Standards. Research Paper, No. RP946 (1936).

126. Gravity and gravitation. We have defined the force of **gravity,** or weight, of a body as the force with which the body is pulled toward the earth as measured by a spring balance. This would clearly include the effect of centrifugal force due to the earth's rotation. **Gravitation,** on the other hand, is the attraction alone of the earth for the body, in accord with Newton's general law, and does not include the effect of centrifugal force.

127. Gravitation inside a homogeneous sphere. It may readily be shown from the law of universal gravitation that inside a homogeneous hollow sphere the gravitational force due to the shell is *zero.*

From this fact, the rather surprising result may be deduced that within a homogeneous solid sphere the force of gravitation varies directly as the distance from the center.

EQUILIBRIUM UNDER GRAVITY

128. Stability. Most structures (houses, automobiles, step-ladders, boats, etc.) depend upon the force of gravity to keep them upright. A body is said to be **stable** when it is not easily overturned. **Stability** is measured by the work required to overturn the body.

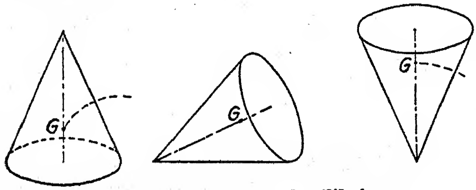


FIG. 113. States of Equilibrium

It is customary to distinguish three states of equilibrium (usually also rest) under the action of gravity (Fig. 113):

1. A body is said to be in **stable equilibrium** when, after any small overturning motion, it tends to return to its former position; e.g., a cone standing on its base. Its **potential energy is a**

minimum, for any overturning motion raises its center of gravity.

2. A body is said to be in **neutral**, or **indifferent, equilibrium** when it remains in any position in which it is placed; e.g., a cone lying on an element of its conical surface. Its **potential energy is the same in all positions,** for the height of its center of gravity remains unchanged.

3. A body is said to be in **unstable equilibrium** when, after any small overturning motion, it tends to move away from its position of equilibrium; e.g., a cone standing on its apex. Its **potential energy is a maximum,** for any motion lowers its center of gravity.

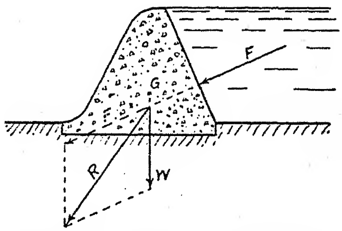


FIG. 114. Criterion of Stability

129. Criterion of stability. A body will be in a condition of stable equilibrium provided the line of action of the resultant of the forces upon it falls within its base, and any motion tends to raise its center of gravity.

A dam, for example, will not overturn so long as the resultant of its weight and the force of the water does not fall beyond the base of the dam. For safety, however, engineers specify that the resultant R must fall *within the middle third* of the base. (See Fig. 114.)

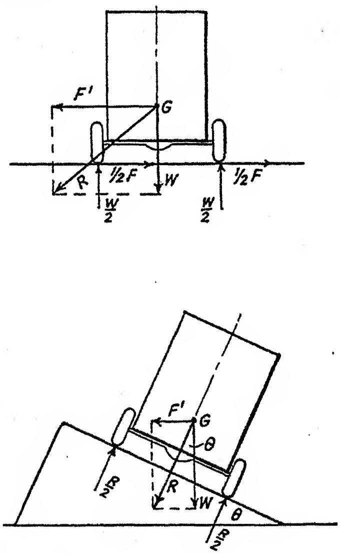


FIG. 115. Banked Curve

130. Banked curves (superelevation). Railroad and highway curves afford other illustrations of this principle of stability.

In order for a car (Fig. 115) to negotiate a curve at legal speed, a centripetal force must be provided sufficient to overcome the tendency of the car, due to its inertia, to continue in a straight line and fly off the curve on a tangent.

If the legal speed for which the road is designed is v and its radius is r, we know that the centripetal acceleration a_0 necessary

to keep the car from flying off the curve is (from Sec. 60),

$$a_0 = \frac{v^2}{r}$$

and the centripetal force F required is, by Newton's second law,

$$F = Ma_0 = M\frac{v^2}{r} \quad \text{(provided absolute or B.E. units are used)}$$

and must be supplied by the friction of the roadbed against the tires.

The inertial reaction, or centrifugal force F', which acts at the center of gravity of the car, is equal and opposite to F, and these form a couple which on a level road tends to overturn the car. In addition, if the friction is not sufficient to keep the car on the curve, there will be skidding.

If the resultant R of F' and the weight W of the car falls outside the base, as is shown in Fig. 115a, the car will overturn; and even for smaller values of F' there is a tendency to smash the wheels.

All these difficulties are avoided if the roadbed is banked, or superelevated, at such an angle θ that R is perpendicular to the roadbed at the midpoint between the wheels, as in Fig. 115b.

In that case, R has no component parallel to the roadbed, and hence there is no tendency to skid or overturn.

To express this condition mathematically, we observe that, since R is perpendicular to the roadbed and W is perpendicular to the horizontal, the angle between R and W equals the angle θ at which the road is banked. Hence, from the figure,

$$\tan \theta = \frac{F'}{W} = \frac{M\frac{v^2}{r}}{Mg} = \frac{v^2}{rg} \tag{134}$$

where it was necessary to express W in absolute units since F' was already so expressed.

At any speed other than that for which the road was designed, there will be some tendency to skid and overturn, but overturning will not take place unless R falls outside the wheel base.

131. Criterion of floating stability. Whether a floating body will right itself or overturn after being tipped from an upright position depends upon the relative positions of its center of grav-

ity G and its metacenter M (Fig. 116). The metacenter is the point where a vertical line through the center of buoyancy B cuts the center line of the vessel. The center of buoyancy B is the center of gravity of the displaced fluid.

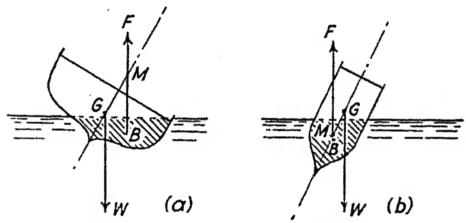


FIG. 116. Floating Stability

If M is above G, as in Fig. 116a, the weight W and the force of buoyancy F constitute a couple tending to right the vessel.

On the other hand, if M is below G, as in Fig. 116b, this couple tends to make the vessel capsize.

Since ballast or a heavy keel lowers G, it increases the stability of the boat.

PROBLEMS

1. At what angle must a railroad curve of 500-ft radius be banked so that the reactions on the rails will be equal for a train speed of 60 mph?

2. If the radius of a curve is 300 ft, at what angle should it be banked in order that a car traveling 54 mph will have no tendency to skid?

3. An automobile having a mass of 2000 lb goes around a curve of 200-ft radius at a speed of 30 mph. Find the centrifugal force. At what angle must the curve be banked so that the car will not tend to skid?

4. What is the greatest speed a car can make without skidding on a level curve of 100-ft radius, if the coefficient of friction of rubber on road is 0.6?

5. A highway curve of 200-ft radius is banked 34° 10′. What is the maximum speed in miles per hour that a car may make without tendency to skid?

6. At what height above the surface of the earth would a body weigh 1/3 as much as at the surface? (Radius of earth = 4000 mi.)

7. What is the acceleration of Martian gravity if the mass of Mars is 1/9 that of the earth and its diameter is 4000 mi, the diameter of the earth being taken as 8000 mi?

8. At what distance above the earth would a body weigh half as much as at the earth's surface, taking the earth's diameter as 8000 mi?

9. If the earth's equatorial radius is 4000 mi, what must its period of rotation be in order that gravity at the equator may be zero?

10. If the moon's mass is 1/80 the mass of the earth and its diameter 1/4 that of the earth, how much would a man weigh on the moon who weighs 150 lb on the earth?

CHAPTER IX

MACHINES

132. A **machine** is any device for transmitting or transforming energy.

However complicated a purely mechanical machine may be, such as a printing press or a harvester, it may be resolved into not more than **six simple, or elementary, machines** (sometimes incorrectly called mechanical powers). These are:

1. Lever
   (a) Wheel and axle.
2. Inclined plane
   (a) Screw
   (b) Wedge.
3. Pulleys.

The lever, the inclined plane, and the pulleys are the *fundamental* machines, the others being special cases of these.

133. Actual mechanical advantage. When we use a machine—say, a claw hammer to draw a nail, or a jack to lift a car—we do so because there is some advantage in that way of doing the job. In order that we may compare different machines as to their usefulness for a given purpose, the term **actual mechanical advantage is defined as the ratio of the useful force F_2 got out of the machine to the force F_1 applied to (or put into) the machine.** Algebraically,

$$\text{Actual mechanical advantage} \equiv \frac{\text{Force out}}{\text{Force in}} \equiv \frac{F_2}{F_1}. \qquad (135)$$

It is important to note carefully the distinction between mechanical advantage and efficiency. By Eq. (100)

$$\text{Efficiency} \equiv \frac{\text{Work out}}{\text{Work in}}.$$

Mechanical advantage may be greater than, equal to, or less than 1; but efficiency is always a fraction less than 1.

134. Ideal mechanical advantage. Since in any actual machine, part of the effort F_1 is always required to overcome friction, actual mechanical advantage is always less than it would be if the machine were **ideal,** i.e., without friction.

Consider such an ideal, or frictionless, machine. Its efficiency would be 100%, or 1. Let F_1 act through the distance s_1 in the direction of F_1; and let F_2 act through the distance s_2 in the direction of F_2.

Then by Eq. (82), Work in $= F_1 \times s_1$

and Work out $= F_2 \times s_2$

and by Eq. (100), Efficiency $\equiv \dfrac{\text{work out}}{\text{work in}} = \dfrac{F_2 s_2}{F_1 s_1}$

whence
$$\frac{F_2}{F_1} = (\text{eff.})\frac{s_1}{s_2}. \tag{135a}$$

But by hypothesis the efficiency is 1.

Therefore **for an ideal machine,**

$$\frac{F_2}{F_1} = \frac{s_1}{s_2}.$$

Consequently **the ideal mechanical advantage of a machine is defined as the ratio of the distance s_1 through which the effort acts to the distance s_2 through which the resistance acts.**

Algebraically,

$$\text{Ideal mechanical advantage} \equiv \frac{s_1}{s_2}. \tag{136}$$

But by Eq. (135), $\dfrac{F_2}{F_1} \equiv$ actual mechanical advantage.

Therefore from Eq. (135a) it will be seen that

Actual mech. adv. = efficiency × ideal mech. adv. (136a)

135. The lever. A lever consists usually of a bar, or rod, having a bearing at some point B, called the **fulcrum** (Fig. 117).

A force called the **effort** F_1 does work on the lever, which in turn does work against the force F_2 called the **resistance,** or

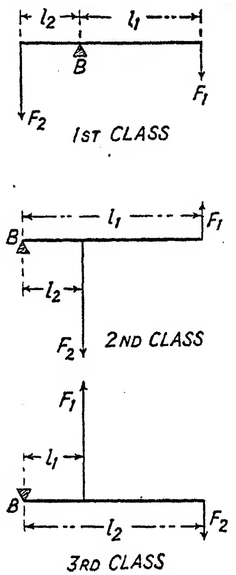


FIG. 117. Three Classes of Levers

load, F_2; and the equilibrant E which acts at B is called the **reaction** of the fulcrum.

Three possible arrangements of effort, fulcrum, and resistance at once suggest themselves, and give rise to the three classes of levers, as shown in Fig. 117.

1. In levers of the first class, the fulcrum lies between F_1 and F_2; e.g., the seesaw.

2. In levers of the second class, F_2 lies between the fulcrum and F_1; e.g., the wheelbarrow.

3. In levers of the third class, F_1 lies between the fulcrum and F_2; e.g., the forearm.

136. Mechanical advantage of the lever. Consider any lever such as the bent lever, or bell crank, of Fig. 118. Let F_1 and F_2 be kept perpendicular to l_1 and l_2, respectively. Then when F_1 turns the lever through the angle θ,

the distance through which F_1 acts is $s_1 = l_1\theta$

and the distance through which F_2 is overcome is $s_2 = l_2\theta$.

Hence

ideal mechanical advantage

$$\equiv \frac{s_1}{s_2} = \frac{l_1\theta}{l_2\theta} = \frac{l_1}{l_2}$$

and

actual mechanical advantage

$$\equiv \frac{F_2}{F_1} = (\text{eff.})\frac{l_1}{l_2}. \quad (137)$$

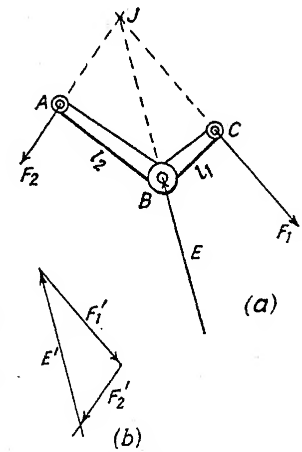


FIG. 118. Bent Lever, or Bell Crank

Hence, **the law of the lever is that the effort and the resistance are inversely proportional to their lever arms, if friction is neglected.**

The same result may be obtained by employing the principle of work.

137. The wheel and axle consists of two cylinders rigidly fastened together so as to turn about the same axis, or bearing, B.

The effort F_1 unwinds a cord from the circumference of the larger, causing the load F_2 to be moved by the other cord winding on to the small cylinder at A. (See Fig. 119.)

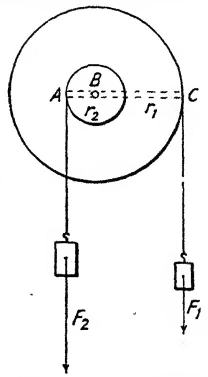


FIG. 119. Wheel and Axle

If we think of a slice through the device as shown by the dotted lines, the remainder of both cylinders being removed, F_1 would still balance F_2, and we should have a lever ABC of the first class.

As the cylinders turn, other strips of the cylinders take up the position ABC, so that the wheel and axle is just a continuously acting lever of the first class. Its mechanical advantage, determined as in the preceding section, is therefore:

$$\text{Actual mechanical advantage} \equiv \frac{F_2}{F_1} = (\text{eff.})\frac{r_1}{r_2}. \tag{138}$$

138. The inclined plane presents two cases.

Case I: When the effort F_1 is parallel to the incline (Fig. 120a).

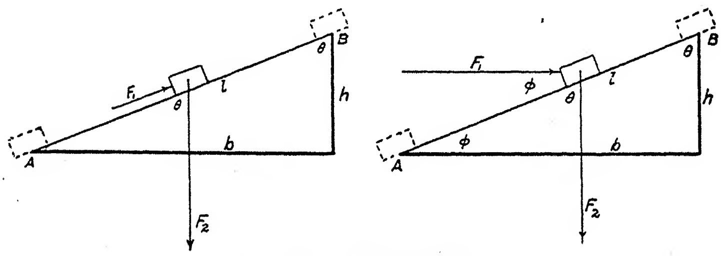


FIG. 120. Inclined Planes

Using the principle of work, when the body is moved from A to B:

$$\begin{aligned} \text{Work put in by } F_1 &= F_1 \times l \times 1 \\ \text{Work out against } F_2 &= F_2 \times l \times \cos\theta \\ &= F_2 \times h. \end{aligned}$$

From Eq. (100)

$$\begin{aligned} \text{Work out} &= \text{efficiency} \times \text{work in} \\ F_2 h &= (\text{eff.})F_1 l \end{aligned}$$

whence $$\text{Actual mechanical advantage} \equiv \frac{F_2}{F_1} = (\text{eff.})\frac{l}{h}. \tag{139}$$

Case II: When the effort (F_1) is parallel to the base of the right triangle of which the incline is the hypotenuse (Fig. 120*b*). In this case,

$$\text{Work in} = F_1 \times l \times \cos\phi = F_1 b$$
$$\text{Work out} = F_2 \times l \times \cos\theta = F_2 h$$

and by Eq. (100),

$$F_2 h = (\text{eff.})\, F_1 b$$

whence

$$\text{Actual mechanical advantage} \equiv \frac{F_2}{F_1} = (\text{eff.})\, \frac{b}{h}. \qquad (140)$$

139. The screw. A screw thread, or helix, is equivalent to an inclined plane wrapped around a cylinder. The **pitch** p of a screw is defined as the distance between corresponding points on consecutive threads. The **lead** l is the axial distance the screw advances through the nut for one complete turn of the screw with respect to the nut. Hence l will equal p, $2p$, $3p$, etc., according as the screw has a single, double, triple, etc., thread.

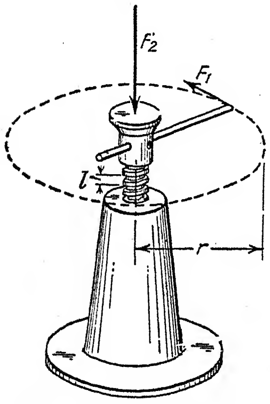


FIG. 121. Jack Screw

To find the mechanical advantage of a screw, consider the jack-screw of Fig. 121. The effort F_1 acts in a horizontal plane and at right angles to a lever arm r. In one complete turn, F_1 acts around the circumference ($= 2\pi r$); and at the same time F_2 is lifted a distance which is the lead l that the screw advances out of the nut. Hence,

$$\text{Work in} = F_1 \times 2\pi r \times 1$$
$$\text{Work out} = F_2 \times l \times 1.$$

By Eq. (100),

$$\text{Work out} = (\text{eff.}) \times \text{work in}$$
$$F_2 l = (\text{eff.})\, F_1(2\pi r)$$

whence

$$\text{Actual mechanical advantage} \equiv \frac{F_2}{F_1} = (\text{eff.})\, \frac{2\pi r}{l}. \qquad (141)$$

The ideal mechanical advantage of screws may be very great, say 1200 or more. However, their efficiency is usually small, since

to be of service, the friction must be so great that the screw is **self-locking;** i.e., it will not run down under the load when the effort is removed.

140. Efficiency of self-locking machines. Consider the case of a body which is just on the point of sliding down an inclined plane (Fig. 122). The force tending to make it slide down is the component of its weight parallel to the plane; that is,

$$F_2 \sin \theta.$$

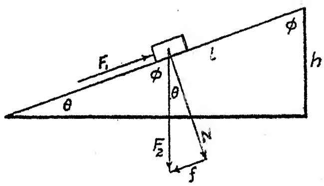


FIG. 122

Since it does not slide, the friction f must be at least as great as that component; hence,

$$f = F_2 \sin \theta.$$

Consequently, when the body is pushed up the plane by F_1,

$$\text{Work put in by } F_1 = F_1 \times l \times 1$$
$$\text{Work done against } F_2 = F_2 \times l \times \cos \phi = F_2 h.$$

If the friction is the same for upward as for downward motion,

$$\text{Work done against } f = f \times l \times 1 = F_2 l \sin \theta = F_2 h.$$

Assuming conservation of energy,

$$F_1 l = F_2 h + F_2 h = 2F_2 h$$

and hence

$$\text{Actual mechanical advantage} \equiv \frac{F_2}{F_1} = \frac{l}{2h}.$$

But by Eq. (139),

$$\text{Ideal mechanical advantage} = \frac{l}{h}$$

and by Eq. (136),

$$\text{Efficiency} = \frac{\text{Actual mechanical advantage}}{\text{Ideal mechanical advantage}}$$

$$= \frac{\frac{l}{2h}}{\frac{l}{h}} = \frac{1}{2} = 50\%.$$

Hence, if friction is the same for upward as for downward motion, the self-locking inclined plane cannot have an efficiency greater than 50%. The same argument may be applied to any type of machine. Wedges, jackscrews, and differential pulleys are self-locking.

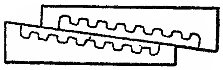


Fig. 123. Printer's Quoins

141. Wedges. The *wedge* is another special case of the inclined plane, and hence its theoretical mechanical advantages would be the same as for the inclined plane. Friction is usually so great in the case of wedges and the direction of the load is so uncertain that a computed mechanical advantage is of little value.

Nevertheless, wedges are of great usefulness in the arts. Often two wedges are used with the two inclined faces together, as in the case of the "quoins" in Fig. 123 with which a printer clamps up his type before placing it in the press.

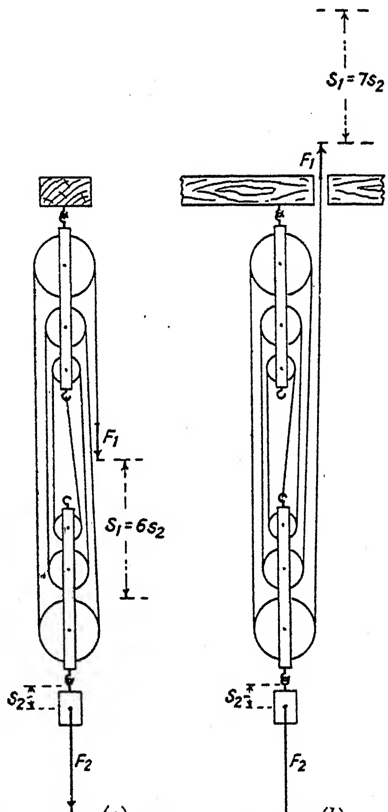


Fig. 124. Pulley Systems

142. Pulleys. A pulley consists of a grooved wheel, or sheave, free to turn on a pin which is fastened in a yoke that carries two hooks, as shown in Fig. 124. The whole device is often called a "block" and may carry any number of wheels.

A **single fixed pulley block**—the wheel, of course, being free to turn—merely changes the direction of the applied force.

When two blocks are used, each containing several sheaves, the mechanical advantage of the "block and tackle" will depend upon whether the fast end of the rope is tied to the one block or to the other. The type of block shown in Fig. 124 is not used

practically, but it is drawn here so that all the rope segments may be seen.

143. Two-block systems of pulleys are best illustrated by two cases.

Case I: When the rope is tied to the hook of the upper block (Fig. 124a). Here it is seen that in order to raise the load a distance s_2, the lower block must be raised a distance s_2; hence all six ropes that support the lower block must be shortened by the amount s_2. The entire amount of this shortening must be taken up by F_1, which must therefore act through a distance $6s_2$. That is,

$$s_1 = 6s_2.$$

Hence, by Eq. (136),

$$\text{Ideal mechanical advantage} \equiv \frac{s_1}{s_2} = \frac{6s_2}{s_2} = 6$$

which is the number of rope segments that support the load.

Case II: When the rope is tied to the hook of the lower block (Fig. 124b).

By the same reasoning as above, we have in this case:

$s_1 = 7s_2$ and hence, by Eq. (136)

$$\text{Ideal mechanical advantage} \equiv \frac{s_1}{s_2} = \frac{7s_2}{s_2} = 7$$

which again is the number of rope segments that support the load.

Hence we may state the general law: **The ideal mechanical advantage of a two-block system of pulleys is equal to the number of rope segments that support the load.**

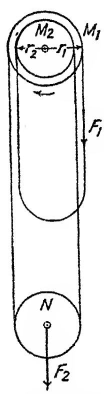


Fig. 125. Differential Pulleys

144. The differential pulley system (Fig. 125) is a combination of a wheel and axle with a movable pulley. The radius r_1 of the wheel M_1 is only slightly greater than the radius r_2 of the axle M_2. M_1 and M_2 are rigidly fastened together and turn on the same shaft. The movable pulley N is carried by a continuous chain whose links fit into depressions in M_1 and M_2 so that there is no slipping.

When the wheel and axle make one complete turn clockwise, a length of chain $2\pi r_1$ winds onto M_1 at the left side, while a length $2\pi r_2$ winds off M_2 at the right side.

The length of the loop carrying N therefore decreases by the amount $2\pi r_1 - 2\pi r_2$; but as N always divides this into two equal parts, the load F_2 is raised only $(2\pi r_1 - 2\pi r_2)/2$. At the same time, F_1 acts through the distance $2\pi r_1$.

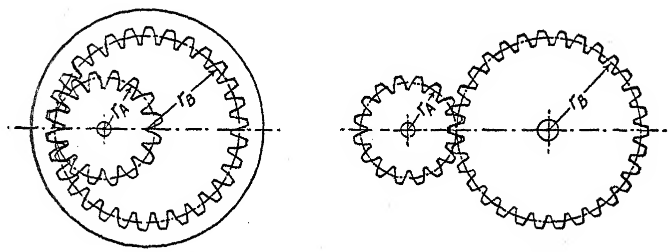
By Eq. (100),

$$\text{Work out} = \text{efficiency} \times \text{work in}$$
$$F_2(\pi r_1 - \pi r_2) = (\text{eff.})\ F_1(2\pi r_1)$$

whence

$$\text{Actual mechanical advantage} \equiv \frac{F_2}{F_1} = (\text{eff.})\ \frac{2r_1}{r_1 - r_2} \cdot \quad (142)$$

145. Toothed gearing. This method of transmitting energy from one part of a machine system to another affords a mechani-



(a) (b)

FIG. 126. Internal and External Gears

cal advantage, while at the same time the teeth assure a **positive drive,** i.e., no slipping.

Consider a pair of toothed gears A and B. They will rotate in the same or in opposite directions according as they engage internally (Fig. 126a) or externally (Fig. 126b).

Let $\mathfrak{T}_A$ be the torque applied by A, and $\mathfrak{T}_B$ the torque delivered by B. Also let N_A and N_B be the rotational speeds of A and B, respectively in rpm.

Then, in 1 minute:

The angle turned through by A is $\theta_A = 2\pi N_A$ rdn
The angle turned through by B is $\theta_B = 2\pi N_B$ rdn

and by Sec. 77,

$$\text{Work done by } A\text{, i.e., work in } = \mathfrak{T}_A\theta_A = \mathfrak{T}_A(2\pi N_A)$$

and

$$\text{Work done by } B\text{, i.e., work out } = \mathfrak{T}_B\theta_B = \mathfrak{T}_B(2\pi N_B).$$

Since

$$\text{Work out} = \text{efficiency} \times \text{work in}$$
$$\mathfrak{T}_B(2\pi N_B) = (\text{eff.})\ \mathfrak{T}_A(2\pi N_A)$$

whence

$$\text{Actual mechanical advantage} \equiv \frac{\mathfrak{T}_B}{\mathfrak{T}_A} = (\text{eff.})\ \frac{N_A}{N_B}. \tag{a}$$

Let T_A and T_B be the number of teeth on A and B, respectively; and let r_A and r_B be the radii of their pitch circles. By *pitch circles* are meant imaginary circles (on the gears) which roll on each other without slipping as the gears turn.

In order that gears may operate smoothly, the width of a tooth of one at the pitch circle must equal the width of the space on the other, and vice versa. Hence, the width of a tooth and a space on A must equal the width of a space and a tooth on B at the pitch circles.

Call the width of a tooth and a space at the pitch circle $\equiv p$.

Then

$$\text{the circumference of pitch circle of } A \text{ is } 2\pi r_A = pT_A \tag{b}$$

and

$$\text{the circumference of pitch circle of } B \text{ is } 2\pi r_B = pT_B. \tag{c}$$

Also, the linear speed of any point on pitch circle of A is

$$V_A = 2\pi r_A N_A = N_A p T_A$$

and the linear speed of any point on pitch circle of B is

$$V_B = 2\pi r_B N_B = N_B p T_B.$$

Since there is no slipping,

$$V_A = V_B.$$

Therefore

$$2\pi r_A N_A = 2\pi r_B N_B \tag{d}$$

and

$$N_A p T_A = N_B p T_B. \tag{e}$$

From Eqs. (a), (d), and (e),

$$\text{Ideal mechanical advantage} = \frac{N_A}{N_B} = \frac{T_B}{T_A} = \frac{r_B}{r_A}. \tag{143}$$

146. Power transmitted by belts. By analysis precisely similar to the preceding, it is found that the mechanical advantage and relative speeds for belted wheels also are given by Eq. (143).

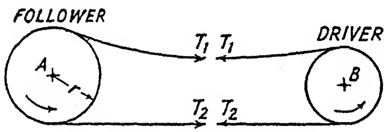


FIG. 127. Belted Wheels

To find an expression for the power transmitted by a belt, let A be the driven pulley in Fig. 127. Also let T_2 and T_1 be the tensions on the tight and loose sides, respectively, of the belt, and let r be the radius of A.

Then, from Eq. (60), the torque on pulley A is:

$$\mathfrak{T} = T_2 r - T_1 r = (T_2 - T_1)r.$$

By Eq. (83), the energy transmitted in 1 sec is:

$$W = \mathfrak{T}\omega = (T_2 - T_1)r\omega.$$

But by Eq. (35), $r\omega = V$, the linear speed of a point on the rim of wheel A, which is also the linear speed of the belt if there is no slipping. And since power P is work, or energy, per unit time,

$$P = (T_2 - T_1)V \tag{144}$$

where P is in ergs/sec when cgs units are used.

If T_2 and T_1 are in lb force, r in ft, and N in rpm, then $V = 2\pi rN$, and Eq. (144) becomes:

$$P = \frac{2\pi rN(T_2 - T_1)}{33{,}000}\ \text{hp}. \tag{145}$$

PROBLEMS

1. How may a 16-ft scantling be used as a lever so that a force of 50 lb may lift a body that weighs 800 lb?

2. If the doubletree on a wagon is 42 in. long, where must it be attached to the wagon in order that a colt may pull half as much as a horse?

3. A hiker carries a bag weighing 8 lb by means of a stick 30 in. long over his shoulder. If it rests on his shoulder at a point 10 in. from the bag, what is the force on his shoulder?

4. A boy capable of exerting a force of 40 lb wishes to get a 200-lb block of ice into a wagon 2.5 ft high by using an inclined plane. What is the shortest board that will serve his purpose if friction is neglected?

5. The great stones of the Pyramids are supposed to have been drawn up on inclined planes. How many men, each capable of exerting 100 lb force, would be required to haul a 9-ton block up an incline 300 ft long and 40 ft high, if by rollers the efficiency were 30%?

6. A weight of 300 lb is to be lifted by a windlass (wheel and axle). If the diameter of the axle is 6 in. and the crank is 15 in. long, what force must be applied at the end of the crank if the efficiency is 60%?

7. A safe weighing 800 lb is to be lifted by a force of 120 lb. What is the smallest number of pulleys by which this may be done (same number in each block)? Sketch the arrangement.

8. Using two blocks having the same number of pulleys, what is the smallest number in each that will enable a piano weighing 650 lb to be lifted by a force of 150 lb? Sketch the system.

9. If the efficiency of a jackscrew is 30%, what is the greatest load that can be lifted by a force of 60 lb if the pitch of the screw is 1/7 in. and the turning bar is 2 ft long? What is the ideal mechanical advantage? What is the actual mechanical advantage?

10. A jackscrew has a pitch of 0.1 in. If the efficiency is 20%, what length of turning rod must be used to enable a force of 50 lb to lift a body weighing 15,000 lb?

11. Two shafts are connected by gears having 20 and 90 teeth respectively. If the speed of the first shaft is 120 rpm, what is the speed of the second shaft?

12. An engine fly-wheel 4 ft in diameter makes 250 rpm and is belt connected to a pulley 2.5 ft in diameter on a line shaft. What is the speed of the line shaft?

13. A table saw must run at 3000 rpm and requires a 3-hp motor which runs at 3600 rpm. If the pulley on the saw mandrel is 4 in. in diameter, what must be the diameter of the pulley on the motor shaft, allowing 8% for slip? If the efficiency of transmission is 90%, what is the difference in the tensions on the two sides of the belt?

CHAPTER X

PENDULUMS

147. The simple pendulum. A casual observation of Galileo's in 1583, while at prayer in the Cathedral of Pisa, is the basis of most of our measurements of time. A large swinging lamp vibrated in what **seemed** to him equal intervals of time. He then designed the first pendulum clock; but the invention of this indispensable accuser is commonly accredited to the Dutch physicist, Christian Huyghens, who fifteen years later "invented" the same device.

It is often said that Galileo discovered the "isochronism" of the pendulum. This is not quite correct, for we do not know yet that a pendulum beats equal intervals of time. What he really did was to **define** as equal the time intervals of successive swings of a pendulum when the amplitude is small. We accept his definition, and computations and conclusions based upon it invariably check with other phenomena such as the period of the earth's rotation or that of an oscillating crystal—but there is no way to prove his assumption rigorously.

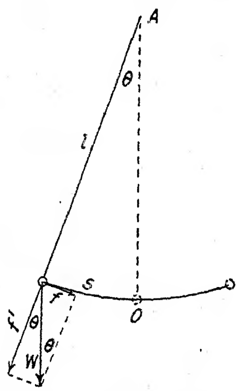


FIG. 128. Simple Pendulum

A simple pendulum is defined as a heavy particle suspended by a weightless thread. A complete vibration is the motion of a pendulum from the time it passes any point, until it again passes through the same point going in the same direction. The **period** is the time required for a complete vibration. The **phase** at any instant is the fraction of a complete vibration that has been completed up to that instant. The **amplitude** is the maximum distance a particle vibrates to either side of its position of equilibrium.

To derive the expression for the period of a simple pendulum, consider a particle of mass of M gm suspended from a fixed axis A by a weightless thread of length l cm (Fig. 128).

Its weight W is then:

$$W = M \text{ gm force}$$
$$= Mg \text{ dynes}$$

and this can be resolved into components f' and f along and perpendicular to l, respectively.

The force f' clearly has no effect except to keep the thread taut; while f produces an acceleration to the right when the displacement (arc s) is to the left, and vice versa.

By Newton's second law,

$$f = Ma. \tag{a}$$

But since θ is clockwise when f is to the right and vice versa,

$$f = -W \sin \theta$$

and for small angles,

$$\sin \theta \doteq \theta \text{ rdn}$$

so that for angles less than 4.5°

$$f \doteq -W\theta.$$

But

$$\theta = \frac{s}{l} \text{ rdn.}$$

Therefore

$$f \doteq -Mg\left(\frac{s}{l}\right). \tag{b}$$

Equating f from Eqs. (a) and (b),

$$Ma \doteq -Mg\frac{s}{l}$$

$$a \doteq -\frac{g}{l}s. \tag{c}$$

Since g and l are constant, Eq. (c) states that the acceleration is proportional to the displacement and in the opposite direction, as is indicated by the $(-)$ sign.

Hence, for very small angles, the simple pendulum has approximately simple harmonic motion (Sec. 35).

For shm,

$$a = -\omega^2 s. \tag{d}$$

Therefore, by Eqs. (c) and (d),

$$-\frac{g}{l}s \doteq -\omega^2 s$$

$$\omega^2 \doteq \frac{g}{l}$$

$$\omega \doteq \sqrt{\frac{g}{l}}.$$

By Eq. (43),
$$T = \frac{2\pi}{\omega}.$$

Therefore,
$$T \doteq \frac{2\pi}{\sqrt{\frac{g}{l}}}$$

$$T \doteq 2\pi\sqrt{\frac{l}{g}}. \tag{146}$$

The result is obviously true for both cgs and fps units.

For amplitudes (i.e., arc s or angle θ) less than 4.5°, the angle expressed in radians equals its sine to one part in a thousand; and Eq. (146) gives T with the corresponding accuracy.

The following more accurate expression,

$$T = 2\pi\sqrt{\frac{l}{g}\left(1 + \frac{1}{4}\sin^2\frac{\theta}{2} + \cdot\cdot\cdot\right)} \tag{147}$$

where θ is the amplitude in radians, permits the effect of amplitude to be taken into account when necessary.*

148. The compound, or physical, pendulum. Since the simple pendulum is an ideal affair that cannot be actually realized, it is necessary to employ other types.

Any extended body, suspended so as to vibrate about a horizontal axis under the action of gravity, is called a **compound,** or **physical, pendulum.**

* For derivation, see A. G. Webster's *Dynamics* (New York, G. E. Stechert & Co.), p. 48.

To compute the period of such a pendulum, let AGP be any such body with axis at A and center of gravity at G (Fig. 129).

The weight acting at G is:

$$W = M \text{ gm force}$$
$$= Mg \text{ dynes}$$

and this produces a clockwise torque $\mathfrak{J}$ such that

$$\mathfrak{J} = -Wx = -Mgh \sin \theta$$

the $-$ indicating that the torque is clockwise when θ is $+$, and vice versa.

By Newton's second law for rotation, Eq. (65),

$$\mathfrak{J} = I\alpha \quad \text{provided absolute or B.E. units are used,}$$

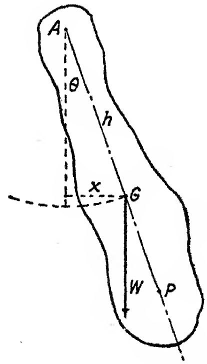


FIG. 129. Compound Pendulum

and for angles less than 4.5°,

$$\sin \theta \doteq \theta \text{ rdn.}$$

Therefore,

$$I\alpha \doteq -Mgh\theta$$

$$\alpha \doteq -\frac{Mgh}{I}\,\theta.$$

Mgh/I is a **constant** for any given pendulum; hence, provided the amplitude is small, the body has approximately **angular shm,** by Eq. (47).

From Sec. 36,

$$c = \omega^2 \doteq \frac{Mgh}{I} \quad \text{for this case}$$

and

$$\omega \doteq \sqrt{\frac{Mgh}{I}}.$$

From Eq. (46),

$$T = \frac{2\pi}{\omega} \doteq \frac{2\pi}{\sqrt{\frac{Mgh}{I}}}$$

so that

$$T \doteq 2\pi\sqrt{\frac{I}{Mgh}} \tag{148}$$

where I is the moment of inertia **about the axis of suspension,** and h is the distance from the axis of suspension to the center of gravity.

The point A is often (somewhat inaccurately) called the **center of suspension.** If the pendulum is reversed, another point P can be found such that if the pendulum oscillates about an axis through P parallel to the former axis through A, the period will be the same as for the axis through A.

The point P is called the **center of oscillation.**

If a body, suspended as in Fig. 129, is struck at different points, there will be a reaction upon the axis due to the blow, different in general for each different point struck. That point at which, if the body is struck, there will be no reaction on the axis due to the blow, is called the **center of percussion.** Obviously a baseball should strike the bat at the center of percussion in order not to jar the batter's hands.

It may be shown both mathematically and experimentally that the center of percussion is the same point as the center of oscillation. The center of suspension, the center of gravity, and the center of oscillation lie in a straight line.

The **equivalent simple pendulum** of a given compound pendulum is the simple pendulum which has the same period as the given compound pendulum.

Equating the periods of the simple and compound pendulums as given by Eqs. (146) and (148), respectively, we have:

$$T \doteq 2\pi\sqrt{\frac{l}{g}} \doteq 2\pi\sqrt{\frac{I}{Mgh}}$$

whence

$$l = \frac{I}{Mh} \tag{149}$$

which is the required **length of the equivalent simple pendulum** for the compound pendulum having I, M, and h.

149. Kater's pendulum. In the deduction of the position of the center of oscillation, the surprising fact comes out that **the distance from the center of suspension to the center of oscillation equals the length of the equivalent simple pendulum.** In 1818 Captain Kater utilized this fact to determine the length of the

equivalent simple pendulum, by attaching to the compound pendulum two knife edges, one of which was adjustable. When adjusted so that the period of the pendulum is the same about both knife edges, the distance between the edges is the length of the equivalent simple pendulum. Such a pendulum is known as Kater's pendulum.

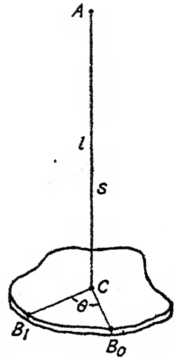


FIG. 130. Torsion Pendulum

150. The torsion pendulum. For obvious reasons, watches and portable clocks cannot use the compound pendulum. In these the time-dividing device is the torsion pendulum.

A torsion pendulum is a body mounted so that it may vibrate about an axis under the action of the torque due to the stress in some attached body, such as the hair-spring of a watch.

Let a body be suspended at any point C by an elastic fiber S from a fixed point A (Fig. 130).

Let a deflecting torque $\mathfrak{T}$ turn the body through an angle θ. Then, provided the elastic limit is not exceeded,

$$\mathfrak{T} = k\theta$$

by Hooke's law (Sec. 175), where k is the **torsional constant.**

As S is twisted, a restoring torque $\mathfrak{T}'$ builds up in it, representing the tendency of the suspending fiber to untwist; and $\mathfrak{T}'$ finally equilibrates $\mathfrak{T}$.

Since there is then equilibrium, we have, by Eq. (76),

$$\Sigma\mathfrak{T} = 0, \quad \text{or} \quad \mathfrak{T} + \mathfrak{T}' = 0$$

and

$$\mathfrak{T}' = -\mathfrak{T} = -k\theta.$$

That is, when the deflection is clockwise, the restoring torque is counterclockwise, and vice versa.

But by Newton's second law, Eq. (65),

$$\mathfrak{T}' = I\alpha$$

Therefore

$$I\alpha = -k\theta$$

$$\alpha = -\frac{k}{I}\theta \qquad \text{(a)}$$

where k/I is constant for a given pendulum.

Comparing Eq. (a) with Eq. (45), we see that the suspended body has angular shm, for which

$$c = \omega^2 = \frac{k}{I}$$

$$\omega = \sqrt{\frac{k}{I}}$$

and by Eq. (43), $$T = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{\frac{k}{I}}} = 2\pi\sqrt{\frac{I}{k}}. \tag{150}$$

One of the best methods of finding the moment of inertia of an irregular body is to determine its period when suspended as a torsion pendulum by a wire of known torsional constant, and then to apply Eq. (150). Great care is required to see that the suspending fiber lies along the axis about which the moment of inertia is required.

The torsion pendulum also gives an excellent method for determining the **modulus of rigidity** η of the material of the suspending fiber. It may be shown that

$$\eta = \frac{2l}{\pi r^4} k * \tag{151}$$

where

η = modulus of rigidity of the material of the wire, or fiber;
l = length of the suspending wire;
r = radius of wire; and
k = torsional constant of wire.

151. Foucault's pendulum experiment. When Copernicus published in 1543 his theory that the earth rotates on its axis, he did so because this gave a much simpler explanation of the apparent motions of the heavenly bodies than was given by the older Ptolemaic theory of a stationary earth and rotating heavens.

The first public experimental demonstration of the earth's rotation was given by Jean Bernard Léon Foucault in 1851.

* For derivation, see Paul L. Bayley and C. C. Bidwell, *Advanced Course in General College Physics* (New York, The Macmillan Company, 1936), p. 30.

From the dome of the Pantheon in Paris, he suspended a ball weighing about 62 lb by means of a wire some 200 ft long. This pendulum was set in motion without jarring, by burning a string

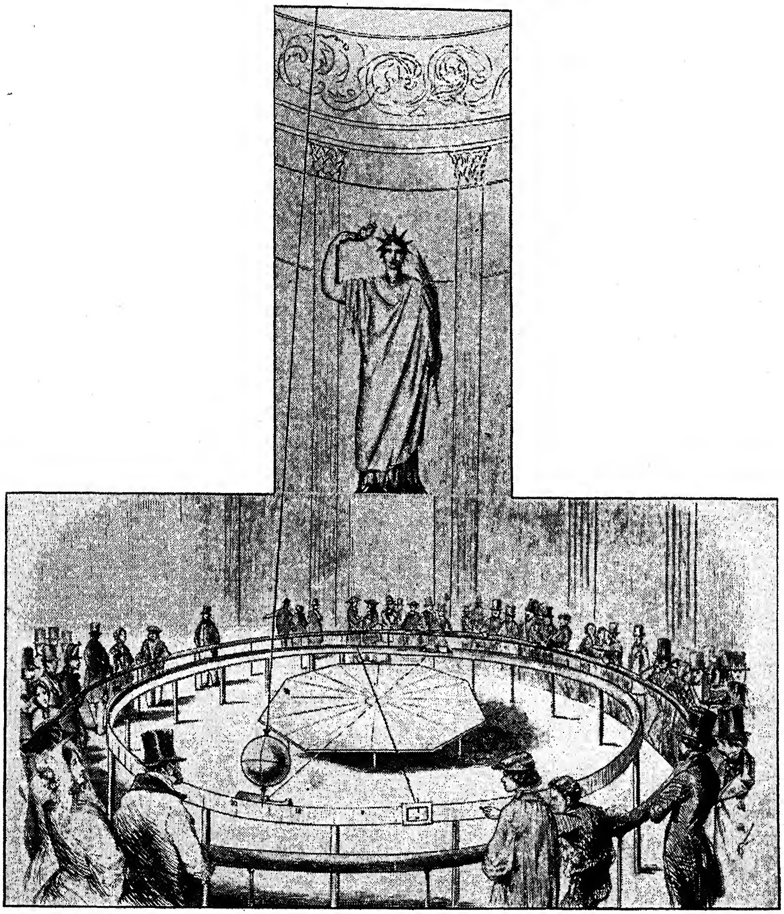


FIG. 131. "Appareil construit par M. Léon Foucault au Panthéon, pour la démonstration du mouvement de rotation de la terre." (From *L'Illustration, Journal Universel*, Vol. XVII, No. 423, April 5, 1851, p. 213.)

with which it had been drawn to one side. Its plane of vibration appeared to rotate clockwise at the rate of about 11° per hour. Actually the pendulum vibrated constantly in the same plane; the apparent rotation was due to the earth's motion under the pendulum.

It may be shown that at any point on the earth's surface the

rate of apparent rotation of the plane of vibration of a pendulum is:

$$\omega = 15 \text{ deg/hr} \times \sin \phi$$

where ϕ is the latitude of the place at which the experiment is made.*

152. Resonance. A body or system is said to make **free vibrations** when its oscillations are unhampered and uncontrolled by any other body or system. When a body or system is making free vibrations, the period is called its **free period;** and the frequency, its **natural frequency.** Thus, if a system consisting of a mass suspended by a helical spring from a rigid support is stretched vertically and released, it will execute free vibrations.

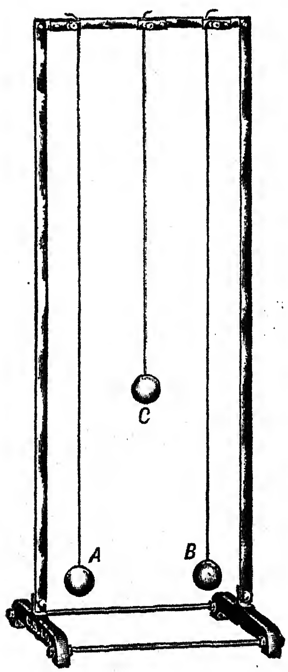


FIG. 132. Resonance Pendulums

When a body is compelled to vibrate with a frequency other than its natural frequency, it is said to make **forced vibrations.** When one pumps water with an ordinary pump (Fig. 136), the vibrations of the pump handle are forced. Again, if the stem of a tuning fork (Fig. 204) is held against a thin table top, the latter is forced into vibration at the frequency of the fork and increases the loudness of the sound but decreases its duration. In the same way, the sounding board of a piano is forced to vibrate at the different frequencies of the various strings.

If two pendulums A and B, having the same free period, or natural frequency, are suspended from a flexible frame (Fig. 132) and A is set swinging, the motion is gradually communicated to B through the flexible frame until B has a **maximum amplitude,** slightly less than the original amplitude of A's motion; and at the same time A comes momentarily to rest. The process is then

* T. W. Wright, *Elements of Mechanics* (New York, D. Van Nostrand Company, 1909), p. 282.

reversed, B giving back the energy to A; and so on until the whole system comes to rest, all of its energy having been dissipated in overcoming the friction of the air and other resistances. Such pendulums are called **resonance pendulums** and are said to be **in resonance.**

Resonance is the phenomenon of the transference of energy from one system to another having the same free period, or natural frequency, as the first system.

If a third pendulum C, having a free period, or natural frequency, *different* from that of A, is suspended from the same flexible frame, it will be set into a slight, halting motion and will absorb very little energy from A. It is not in resonance with A. Its vibrations are the result of its own free vibration and the vibration forced by A; it exhibits what later we shall call **beats.**

Resonance is often observed in nature. It is exemplified in the ringing of a church bell, where the frequency of the pulls on the rope must be the same as the natural frequency of the bell as a pendulum. A bridge may be set into vibrations of dangerous magnitude if its natural frequency is the same as that of the small, regular impulses of marching soldiers. One must employ resonance in rocking a boat, but the practice is strongly discouraged. Tuning a radio receiver consists in getting it in resonance with the sending station. Resonance is also found in the phenomena of sound and light.

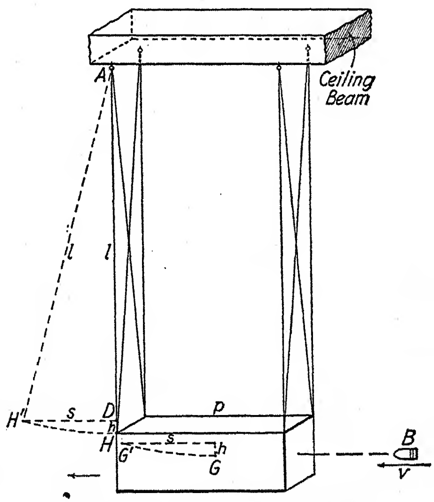


FIG. 133. Ballistic Pendulum

153. The ballistic pendulum. This device, which is much used in measurements involving impact, consists of a heavy body P of mass M gm suspended as in Fig. 133 in order to prevent lateral motion. Its use may be illustrated by showing how it serves to determine the velocity of a bullet.

Let the bullet B of mass m gm be fired so as to strike the pendulum in line with its center of oscillation with the velocity v cm/sec. Both pendulum and bullet then move with a speed V cm/sec, the center of mass G describing an arc like HH' with horizontal displacement s cm.

Consider the system consisting of the bullet and the pendulum.

$$\begin{aligned}
\text{Momentum of bullet before impact} &= mv \text{ gm-cm/sec} \\
\text{Momentum of pendulum before impact} &= \underline{0} \\
\text{Total momentum of system before impact} &= mv \text{ gm-cm/sec}
\end{aligned}$$

$$\begin{aligned}
\text{Momentum of bullet after impact} &= mV \text{ gm-cm/sec} \\
\text{Momentum of pendulum after impact} &= \underline{MV} \text{ gm-cm/sec} \\
\text{Total momentum of system after impact} &= mV + MV \text{ gm-cm/sec.}
\end{aligned}$$

By the law of conservation of momentum,* the total momentum before impact must equal the total momentum after impact. Therefore

$$mv = (m + M)V. \tag{a}$$

To find the value of V, we note that the **work** done in raising the pendulum and bullet the height h is, by Eq. (89),

$$\text{Work} = \text{Potential energy} = (m + M)\,g\,h \text{ ergs} \tag{b}$$

and the **kinetic energy** of the system after impact was, by Eq. (86),

$$\text{K.E.} = \tfrac{1}{2}(m + M)V^2 \text{ ergs.} \tag{c}$$

By the law of conservation of energy,† the potential energy at the top of swing must equal the kinetic energy at the bottom of the swing. Therefore,

$$(m + M)gh = \tfrac{1}{2}(m + M)V^2$$

$$V = \sqrt{2gh}. \tag{d}$$

From the geometry of the figure,

$$h = AH - AD = l - \sqrt{l^2 - s^2}. \tag{e}$$

* Energy is not conserved at this stage, for it is never conserved when heat is generated.

† Momentum is not conserved at this stage. Momentum is never conserved when an external force acts on a system—in this case the pull of the strings and the force of gravity.

Substituting h in Eq. (d) and then V in Eq. (a), we get:

$$v = \frac{m + M}{m} [2g(l - \sqrt{l^2 - s^2})]^{\frac{1}{2}}. \qquad (152)$$

PROBLEMS

1. Compute the length of a seconds pendulum (period = 2 sec) for standard conditions.

2. Compute the length of a simple pendulum to make a complete vibration in 1/2 sec.

3. Find the period of a meter-stick suspended at one end as a compound pendulum. Find the length of its equivalent simple pendulum.

4. A thin hoop 2 ft in diameter is hung over a pin so that it vibrates in its own plane as a pendulum. Find its period.

5. A pendulum consists of a sphere 5 cm in diameter and having a mass of 200 gm suspended by a wire 97.5 cm long, 0.1 mm in diameter, and weighing 1.2 gm. Compute the periods of the pendulum with and without taking the wire into account.

6. A clock pendulum consists of a rod 90 cm long having a mass of 500 gm, with its lower end attached at the center of a flat disk 20 cm in diameter and weighing 1 kg. Compute the period of the pendulum when suspended at the top of the rod and vibrating in the plane of the disk.

7. A rectangular block of wood 8 × 12 × 3 cm has a mass of 300 gm and is suspended by a wire at the center of its 8 × 12 face so as to form a torsion pendulum. If its period is 2.4 sec, what is the torsional constant of the wire?

8. If the wire in the above problem is 0.2 mm in diameter and 30 cm long, what is its modulus of shear?

CHAPTER XI

HYDRAULIC MACHINES

154. The hydraulic press consists of a large cylinder C (Fig. 134), into which water or oil is pumped by means of a small piston from a tank T. The downward force f is distributed over the area a of the small piston. This produces on the enclosed liquid a pressure,

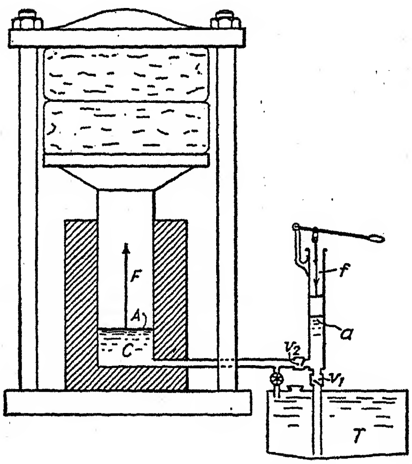


FIG. 134. Hydraulic Press

$$p = \frac{f}{a}$$

which, by Pascal's principle, is transmitted to each unit area of the large piston. The total upward force F of the large piston is therefore:

$$F = pA = \frac{f}{a}A$$

$$\frac{F}{f} = \frac{A}{a}. \tag{153}$$

With slight modifications this device is used as the hydraulic lift, the hydraulic jack, and for barbers' and dentists' chairs.

155. The siphon is a device for conveying a liquid from a higher to a lower level over an intervening obstacle by means of atmospheric pressure and the force of gravity. It usually consists of a pipe, or tube, bent to the required shape, the shape having no effect upon the theory of its action.

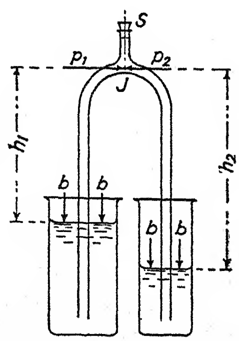


FIG. 135. Siphon

Consider a particle J (Fig. 135). It will move to the right or left according as the pressure to the

right p_1 is greater or less than the pressure to the left p_2. By Eq. (113),

$$p_1 = b - h_1 D$$
$$p_2 = b - h_2 D$$

where b is the barometric pressure and D is the density of the liquid.

Since

$$h_2 > h_1,$$
$$b - h_1 D > b - h_2 D \quad \text{or} \quad p_1 > p_2$$

the particle J and all the other particles will flow to the right; i.e., from the higher water surface to the lower. The direction of flow may be altered at will by raising either beaker and lowering the other.

The atmospheric pressure on the higher surface of the liquid is necessary to keep the tube filled. An air leak in the tube, unless extremely small, will admit sufficient air to counterbalance b and stop the flow, as may be seen by loosening the stopper s.

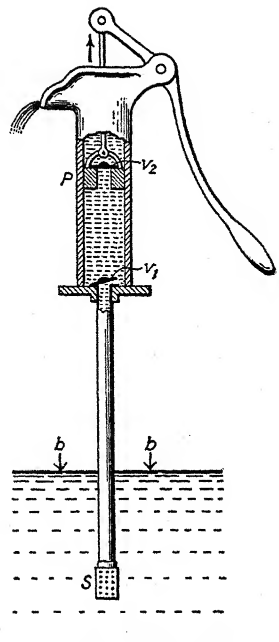


FIG. 136. Lift Pump

156. Lift and force pumps. The two most common types of pumps for raising liquids differ chiefly in the location of the valve v_2. In each, when the piston P is raised, a partial vacuum is created beneath it in the cylinder. **Atmospheric pressure** b then forces water through the strainer s, up the pipe, and through the "foot valve" v_1 into the cylinder.

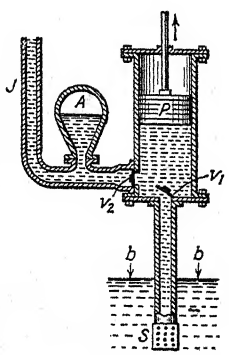


FIG. 137. Force Pump

The **lift pump** (Fig. 136) has the valve v_2 in its piston. On the down stroke v_1 closes and some of the water passes through v_2 and is "lifted" out on the next up stroke.

The **force pump** (Fig. 137) has a solid piston and its valve v_2 is in the delivery pipe. On the down stroke v_1 closes and part of the

water is "forced" up the delivery pipe J. Since atmospheric pressure is equivalent to the pressure of a column of water about 34 ft high, the cylinder cannot be at a height greater than 34 ft above the water surface in the reservoir.

On account of air leakage at the piston packing and elsewhere, and vapor pressure (Sec. 287) above the water column, *actual pumps must have the cylinder within 28 or 30 ft of the water surface.*

After the water enters the cylinder, it may be lifted or forced to any desired height.

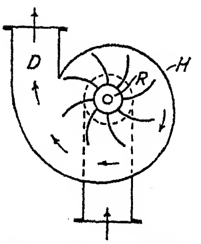


Fig. 138. Centrifugal Pump

157. The centrifugal pump. Because of the ease of connecting them to motors and of the absence of valves, centrifugal pumps have come into general use not only as pumps for ordinary liquids, but also as blowers and exhaust fans for blast furnaces, planing mills, vacuum cleaners, etc.

These pumps consist of a rotor R (Fig. 138) having vanes, radial or shaped as shown, and rotating within a housing H. Particles within the housing are swept forward by the vanes and acquire a high circumferential speed.

On account of "centrifugal force" (actually, the absence of centripetal force necessary to restrain them), these particles fly off according to Newton's first law into the tangential discharge pipe D. This raises the pressure in D and leaves a partial vacuum around the axis of the rotor. The intake pipe is accordingly connected at the center as shown by the dotted lines.

Centrifugal pumps will operate against back pressures as high as 25 lb/in^2, but by using several in tandem they may be used against any ordinary head. They are particularly adapted for moving large quantities of fluid against comparatively low back pressures, as in ventilating systems and wind tunnels.

158. Air Pumps. All pumps are primarily air pumps, for air must be pumped out before another fluid can enter. Thus, the ordinary hand pump at a well first pumps the air out of its cylinder and the intake pipe. This reduces the pressure in these, and atmospheric pressure on the surface of the liquid outside forces it into the pump.

When used to pump air into a vessel, i.e., to increase the pres-

sure above atmospheric, an air pump is often called an **air compressor.** The ordinary bicycle or automobile tire pump is an example.

If used to pump air **out of** a vessel, i.e., to reduce the pressure below atmospheric, an air pump is called a **vacuum pump.**

159. Rotary, oil-sealed vacuum pump. Pumps of the modern type shown in Fig. 139 are very rapid and produce vacuums as high as 0.0001 mm of mercury. The rotor *A* is mounted eccentrically on its shaft. As it turns, the space *H* behind it increases in volume, while the space *J* ahead of it decreases. Hence the air ahead of it is compressed until its pressure is sufficient to lift the ball outlet valve *L*, and the air is expelled through the exhaust opening *F*. When the point of contact *G* between the rotor and its housing *B* passes the intake port *E*, a new volume of air is entrapped ahead of the rotor and is compressed and expelled as before. The ward *C* is held against the face of the rotor at *K* by the spring *S* and bell-crank *D*, and prevents leakage back from the high-pressure to the low-pressure side. The vessel to be exhausted is connected at *V*.

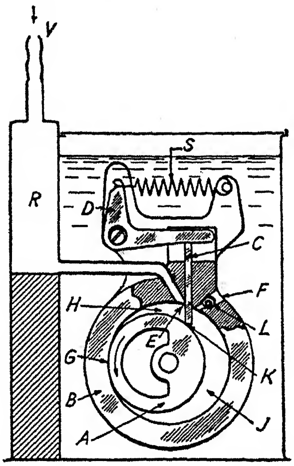


FIG. 139a. Oil Sealed Vacuum Pump (Cross-section)

The entire pumping mechanism is submerged in oil, which effectually seals it against air leaking in. The reservoir *R* has sufficient capacity to hold the oil which is pushed back into the low-pressure side when the pump is shut down, and thereby prevents it from passing into the exhausted system.

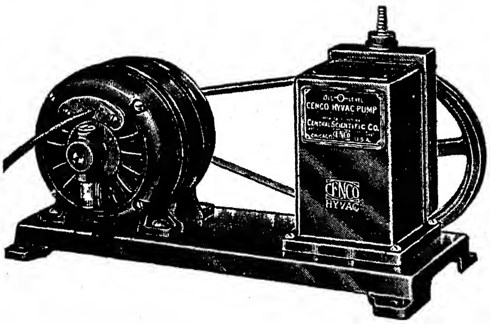


FIG. 139b. Cenco Hyvac Pump. (Courtesy Central Scientific Co.)

Figure 139b shows the pump mounted on a base with its driving motor.

Electric light bulbs are usually evacuated by pumps of this general type. By making connection to the high-pressure side at *F*, this same mechanism is used as a compressor in one type of household refrigerator.

160. Diffusion, or condensation, pumps. The diffusion pump was invented by Gaede, but it is commonly called the Langmuir pump after Dr. Irving Langmuir of the General Electric Company, who perfected it in this country.

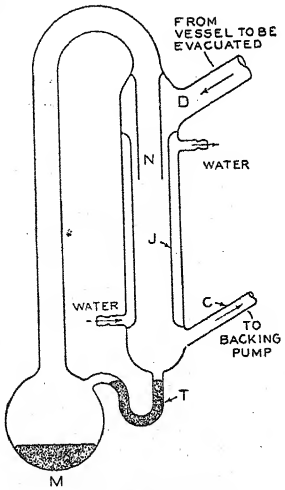


Fig. 140. Diffusion Pump

Originally using mercury, the diffusion pump has been found to work as well or better with certain organic liquids such as butyl phthalate, and with certain oils which have very low vapor pressures. A simple form is shown in Fig. 140. The pump is made either of glass or of steel.

Heat applied at the base *M* vaporizes the liquid and a jet of its molecules rushes out of the nozzle *N*, driving before it the air molecules that diffuse into the tube *J* from the intake *D*.

The continuous jet of mercury molecules serves as an effective barrier against the return of the molecules of air toward *D* and drives them into the outlet tube *C* with such momentum that the pressure there rises to about 0.01 mm of mercury, at which pressure the air can be removed by a backing pump, usually of the oil-sealed rotary type. Some form of backing pump is necessary.

The mercury vapor for the most part condenses on the inner surface of the long tube *J*, and the liquid trickles back into the boiler through the trap *T*. The tube *J* is water-jacketed, as shown.

Some vapor, however, diffuses back into the system being exhausted and must be removed by a liquid air trap, if objectionable. When properly proportioned this pump is very rapid. With the necessary charcoal and liquid air traps and "getters," it is used for

producing the highest attainable vacuums, such as are required in X-ray tubes, which are of the order of 10^{-7} mm of mercury.

HYDRAULIC MOTORS

161. Water motors. The water motor was one of the earliest means of harnessing the available energy of nature. The older

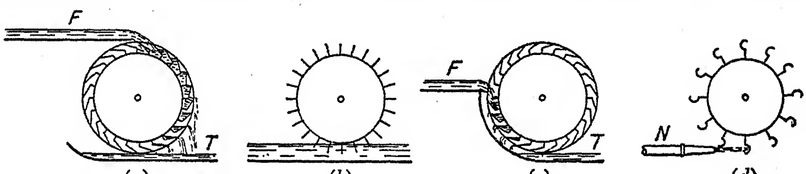


Fig. 141. Water Wheels

forms are illustrated in Fig. 141, namely:

(a) The overshot wheel.
(b) The undershot wheel.
(c) The breast wheel.
(d) The Pelton wheel.

Overshot and breast wheels are driven largely by the dead weight of the water in their buckets. They have been made as large as 100 ft in diameter, and when built of steel and correctly designed, their efficiency may reach 75%.

The undershot wheel is driven by the impact of a rapid stream against its vanes, and the efficiency is very low.

The Pelton impulse wheel employs a nozzle N and a special form of bucket whereby the water is returned parallel to its original direction, and leaves the wheel with very little velocity and hence with little kinetic energy.

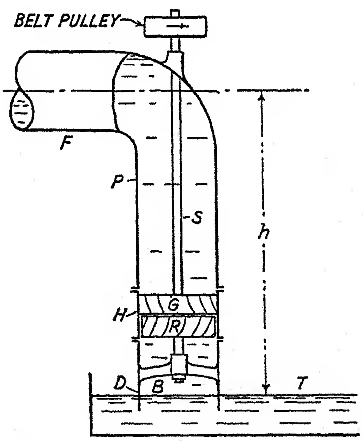


Fig. 142. Turbine

162. The turbine is the modern type of water wheel. Though built with many modifications of details, it consists essentially of a system of "guide vanes" G, fastened rigidly to the housing H (Fig. 142). The "moving vanes" of a second system are part of

the wheel, or "runner," R and are curved opposite to the guide vanes. The runner is keyed to the vertical shaft S which transmits the energy to the driven machinery. The lower end of the shaft runs on a "step-bearing" B.

The water enters through the "penstock" P, passes through the turbine parallel to the shaft, and discharges through the draft tube D into the tailrace T.

Turbines are made in sizes from 10-hp up to the enormous 20,000-hp wheels of the Muscle Shoals plant. Their efficiencies are as high as 75% to 90%.

163. Water power. In Fig. 142, the potential energy of the water in the flume F with respect to the water in the tailrace T is all that might possibly be transformed into useful work by any type of water motor.

Hence the total **energy input** received by the water wheel per minute is, by Eq. (89),

$$W = Mh \text{ ft-lb}$$

where M is the number of pounds of water per minute that flow through or past the wheel, and h is the average difference of level between the water in the flume and that in the tailrace.

Therefore,

$$\begin{aligned} P \equiv \frac{W}{t} &= Mh \text{ ft-lb/min} \\ &= DVh \text{ ft-lb/min} \qquad \text{by Eq. (108)} \end{aligned}$$

and, since $D = 62.5 \text{ lb/ft}^3$ for water,

$$\begin{aligned} P &= 62.5\ Vh \text{ ft-lb/min} \\ &= \frac{62.5\ Vh}{33{,}000} \text{ hp} \end{aligned} \tag{154}$$

where V is the volume of water available per minute in cubic feet.

The useful **output power** may then be obtained by multiplying the input power of Eq. (154) by the **efficiency** of the wheel.

More generally, **for any fluid,** we may find the **total input work** applied to a motor as follows.

Let C (Fig. 143) be a cylinder fitted with a frictionless piston J whose area is A ft^2; and let J be moved from J_1 to J_2, a distance s ft, by a fluid supplied at a constant pressure p lbs/ft^2.

Then the total force acting on the piston is:

$$F = pA \text{ lb force}$$

and the work done on the piston is, by Eq. (82):

$$W = Fs \cos 0° = Fs \text{ ft-lb}$$
$$= (pA)s.$$

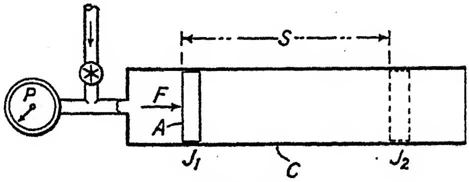


FIG. 143

But As is the volume V ft³ of the fluid required to move the piston from J_1 to J_2. Hence,

$$W = pV \text{ ft-lb} \tag{155}$$

and the input power is

$$P \equiv \frac{W}{t} = \frac{pV}{t} \text{ ft-lb/sec} \tag{156}$$

where t sec is the time required for the volume V of fluid to enter the cylinder.

Since frictionless conditions were assumed, Eq. (155) represents the total energy supplied to the machine by the fluid.

PROBLEMS

1. The diameter of the small piston of a hydraulic jack is 0.5 in., and that of the large piston 1.5 in. What load can be lifted by a force of 50 lb if the mechanical advantage of the lever is 8 and the efficiency 40%?

2. The hydraulic jack has a small piston 0.4 in. in diameter and a large piston 2 in. in diameter. If the efficiency is 60% and the mechanical advantage of the pump lever is 9, what weight can be lifted by a force of 40 lb?

3. On a hydraulic press the diameter of a small piston is 0.5 in. If the force applied to this piston is 100 lb, what must be the diameter of the large piston to enable the press to exert a total force of 20 tons?

4. Assuming an efficiency of 45%, what is the maximum load that can be lifted by a force of 100 lb with a jackscrew of 0.1-in. pitch when a turning bar 2 ft long is used? What is the theoretical mechanical advantage? What is the actual mechanical advantage?

5. Design a hydraulic jack to lift 15 tons, the small piston being 0.5 in. in diameter and having a stroke of 1 in. If the handle gives a mechanical advantage of 6 and the force exerted by the hand is 30 lb, how many strokes will be required to raise the load 4 in.?

6. A stream of water 40 ft wide and 6 ft deep flows at an average speed of 2 mph. If the effective head is 8 ft, what power can be developed by a turbine whose efficiency is 70%?

7. If the cross section of a stream is 30 ft wide by 8 ft deep and its average speed is 2 mph, what power can be developed in using a turbine of 75% efficiency, the effective head being 10 ft?

8. The mean speed of a stream is 1.5 mph. If its cross section is 20 ft wide and 3 ft deep, what power can be developed with an available head of 40 ft and a turbine whose efficiency is 80%?

9. A centrifugal blower delivers 1500 ft^3 of air per minute at a gauge pressure of 0.5 lb/in^2. If the efficiency of the blower is 80%, what must be the power of the motor that drives it?

10. A certain industry requires 229 ft^3 of water per minute. If the water is delivered at a pressure of 50 lb/in^2 and the efficiency of pump and pipe line is 75%, what power is required to drive the pump?

CHAPTER XII

PRESSURE GAUGES AND DYNAMOMETERS

164. The pressure gauges on steam boilers, pressure cookers, gas tanks, etc., are familiar objects. These are commonly called "pressure" or "vacuum" gauges according as they indicate pressures above or below one standard atmosphere. They are calibrated to read **force per unit area**, or the equivalent. Since several gauges depend upon the important principle known as Boyle's law, we take that up at once.

165. Boyle's law. Using the excellent method of measuring pressures which Torricelli had devised for his barometer experiment, Robert Boyle,* thirteen years later (1660), discovered how the volume V of a given mass of a gas changes when its pressure P is altered, the temperature being kept constant.

His method was substantially as follows. In a J-shaped glass tube (Fig. 144) the mercury is adjusted at the bend until the tops of the two columns are tangent to a horizontal line XY.

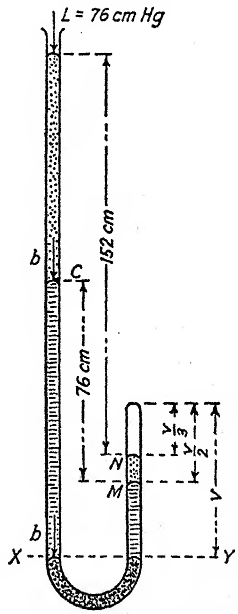


Fig. 144. Boyle's Law

Since the columns are then the same height, they balance each other, and the pressure b of the atmosphere on the mercury in the open tube is transmitted, according to Pascal's principle (Sec. 100), to the air trapped in the closed tube. The pressure on the confined air is then one atmosphere and its volume is V.

* Robert Boyle (1627–1691) was a wealthy Irish gentleman who took up his residence at Oxford University and carried on his experiments privately. Robert Hooke was his assistant.

Mercury is then added until the surface C stands 76 cm above M. The pressure on the confined air has been doubled and the air is found to be compressed to one-half its former volume.

Similarly, when the pressure is trebled the volume is reduced to one-third its original value, and so on.

Tabulating these results:

Pressure	*Volume*
b	V
$2b$	$\frac{1}{2}V$
$3b$	$\frac{1}{3}V$

He therefore stated the law: **At a constant temperature, the volume V of a given mass of gas varies inversely as its pressure P.**

Stated algebraically,

$$V \propto \frac{1}{P} \quad \text{if the temperature is constant,}$$

$$V = C\frac{1}{P}$$

$$PV = C \tag{157}$$

where the value of the constant C will depend upon the mass of gas used, its temperature, and the units in which P and V are expressed.

Boyle extended his mercury columns as high as 117 9/16 inches, the most convenient place for the experiments being on "a pair of stairs." He found that the law is **not exact.** It is sufficiently accurate, however, for most purposes. In general, gases are more compressible than the law requires at moderate pressures, and not compressible enough at high pressures.

166. The open-tube manometer. This type of pressure gauge is used generally for pressures from 0.1 cm to 76 cm of mercury (Hg). It is connected at R to the vessel in which the pressure is to be measured, and at J it is open to the air.

The tube is partially filled with a liquid M. For most purposes mercury is best because it does not wet the tube, but for very low pressures sulphuric acid or colored water is used.

When the surfaces of the liquid in the two sides of the tube are level, the pressure in R is obviously equal to the atmospheric

pressure b as indicated by a barometer. In all cases, the columns below the dotted line XY balance each other, whether or not their areas are equal (by Sec. 99).

When the liquid stands h cm higher on the side open to the atmosphere (Fig. 145a), the pressure is h cm of Hg, say, above atmospheric pressure, for p has to sustain the column h and the pressure b above it. On the other hand, if the liquid stands higher on the side connected to R (Fig. 145b), the atmosphere has to balance the column h and the pressure p; hence the pressure p in the vessel is h cm of Hg less than atmospheric pressure.

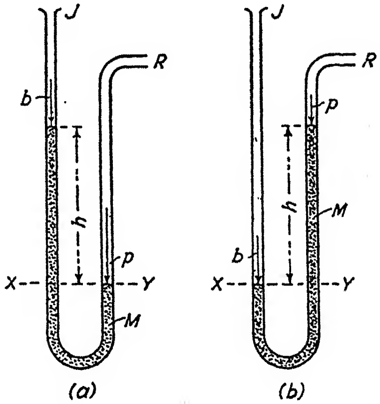


FIG. 145. Open Manometer

In both cases, the total or **absolute pressure p** is expressed by the relation:

$$p = b \pm h \text{ cm of Hg} \quad (158)$$

where the $+$ or $-$ is to be used according as the mercury on the side open to the atmosphere is higher or lower than on the side connected to R.

If a liquid other than mercury is used, let its density be D. Then the height of water equivalent to h cm of the liquid is

$$hD \text{ cm of water.}$$

And since the density of Hg is 13.6 times that of water, the same pressure is produced by

$$h \frac{D}{13.6} \text{ cm of Hg.}$$

Hence, in general, using any liquid of density D,

$$\textbf{Absolute pressure, } p = b \pm h \frac{D}{13.6} \text{ cm of Hg.} \quad (159)$$

167. The absolute gauge. In the foregoing, it was seen that to find the total, or absolute pressure, one must add the barometric height to the reading of an open-tube manometer.

The necessity of reading the barometer is avoided by using an

absolute gauge. This closed-tube manometer (Fig. 146) is inverted and filled with mercury, which is then boiled in the tube to drive out the occluded air and water. On setting the gauge erect again, the closed tube remains completely filled with mercury, which is sustained by atmospheric pressure and has its free surface at X.

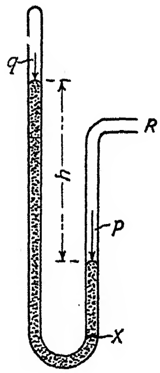


FIG. 146. Absolute Gauge

When the pressure in the vessel connected at R is sufficiently reduced, the mercury surface at the left descends, and the

Absolute pressure, $p = h + q$ cm of Hg (160)

where h is the difference in height of the mercury columns, and q is the vapor pressure of Hg at the temperature at which the experiment is made. At 20°C, $q = 0.001$ mm of Hg, which is less than the usual error of reading the scale; but at higher temperatures q may not be negligible.

168. The Bourdon gauge. For most engineering purposes where an accuracy of 0.1 lb/in.^{2} is sufficient, the Bourdon gauge is

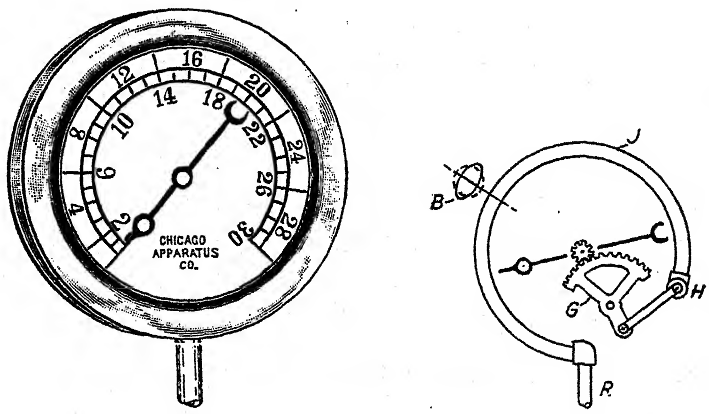


FIG. 147. Bourdon Gauge. (Courtesy Chicago Apparatus Co.)

used. It may be calibrated as either a pressure or a vacuum gauge.

The action depends upon the well-known fact that when the pressure is increased within any body, that body tends to assume a spherical form, because for a given surface the sphere has the

maximum volume. This is seen when one blows up a paper bag, or an elongated toy balloon.

In the Bourdon gauge (Fig. 147b), the tube J of thin metal has an oval cross section B. When pressure is admitted to this tube from the reservoir connected at R, the tube tends to take a circular cross section (dotted), and to straighten out somewhat.

The end H of the tube is connected by links to the segmental gear G, which in turn engages a small pinion on the axis carrying the hand. As the tube straightens out, the point H moves away from the center and the hand moves clockwise over the scale.

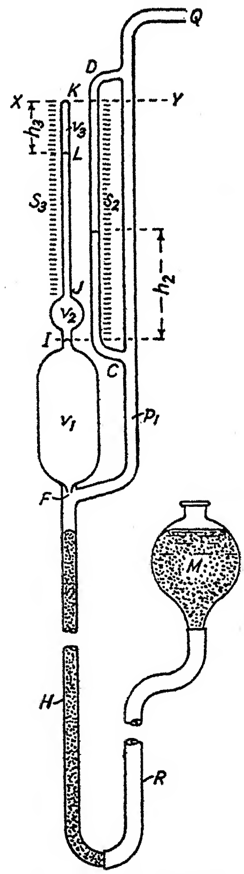


FIG. 148. McLeod Gauge

169. The McLeod gauge. This gauge, invented by H. McLeod in 1874, is used more than any other device for measuring pressures of about 1.5 mm to 0.0001 mm of mercury. The small bulb was added recently by A. H. Pfund and others.

The vessel to be exhausted is connected to the gauge at Q (Fig. 148). The lower glass tube H must be about 85 cm long, and at its lower end it connects by a long rubber tube R to the reservoir of mercury M.

To take a reading, M is raised and mercury flows into bulb v_1, entrapping the gas above the point F.

Case I. Pressures of about 1.5 mm to 0.25 mm are read on scale S_2. The mercury is raised until the top of the left column is at the index I, and the height h_2 at which the mercury stands in the capillary CD is read on scale S_2. Remembering that the pressure at Q is p_1, we then have:

$$p_1v_1 = (h_2 + p_1)v_2$$

by Boyle's law, since $(h_2 + p_1)$ is the final pressure in cm of Hg, and v_2 is the final volume.

Hence,

$$p_1 = \frac{v_2}{v_1 - v_2} h_2. \tag{161}$$

The volumes v_1 and v_2 are accurately determined before the bulb is attached to the rest of the apparatus. Hence the scale S_2 may be laid off to read directly in values of p_1, the pressure in the vessel being exhausted.

Case II. For measuring very low pressures, the mercury is raised until the surface of the column in CD stands at the fiducial line XY. This compresses the entire volume v_1 into the volume v_3 of the capillary LK, the final pressure being h_3 cm of Hg. We now have:

$$p_1v_1 = h_3v_3$$

by Boyle's law as before.
But,

$$v_3 = Ah_3$$

where A is the cross-sectional area of the capillary, which was determined before assembling. Therefore,

$$p_1v_1 = h_3(Ah_3) = Ah_3^2$$

$$p_1 = \frac{A}{v_1} h_3^2. \tag{162}$$

From this relation the value of p_1 for any value of h_3 is computed and the scale S_3 may be laid off to read directly in values of p_1, the pressure in the vacuum system being exhausted.

POWER DYNAMOMETERS

The term **dynamometer** is applied to devices for measuring force, but more generally to those used for the measurement of power.

170. The Prony brake. The power output of an engine or motor is commonly measured by means of a Prony or some other form of friction brake. The Prony brake (Fig. 149) consists of two shoes (usually of wood), notched for ventilation, which are pressed against a pulley K on the engine shaft S by the bolts B,B, until the engine runs at the desired speed. One shoe is extended to

make the brake arm A, which bears upon a knife edge E, which in turn is supported by some form of beam or spring balance.

The drag of the pulley on the brake shoes tends to turn the whole brake clockwise, but this torque is balanced by the counter-clockwise torque of F on the brake arm.

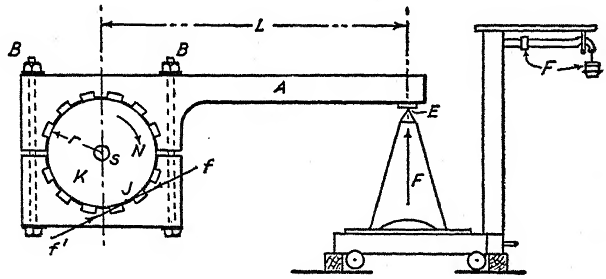


FIG. 149. Prony Brake

Let r cm be the radius of the brake pulley; N rpm be the angular speed of the pulley; f' gm force be the **friction** of the brake on the pulley; f gm force be the **tangential** force of pulley on brake; F gm force be the reaction of the spring or beam balance; and L cm be the lever arm of F.

Since the total friction is independent of the area of contact, we may think of the friction f' of the brake on the pulley as taking place only over the area of a small projection J on the brake. In one revolution, the force f overcomes the friction f' all the way around the circumference $(2\pi r)$. Hence, the work W_1 done by the pulley in overcoming friction in 1 **revolution** is:

$$W_1 = f(2\pi r) \cos \theta. \qquad \text{by Eq. (82)}$$

But f is along the circumference, hence $\theta = 0$ and $\cos \theta = 1$.

Therefore, the total work W_N done by the pulley in **1 minute** is:

$$W_N = Nf(2\pi r) \text{ gm-cm.}$$

Hence the *power* * P absorbed by the brake is:

$$P = 2\pi Nfr \text{ gm-cm/min.} \qquad \text{(a)}$$

Since the brake has no angular acceleration, it is in equilibrium.

* Speaking accurately, it is *energy* that is "absorbed," i.e., transformed into heat, by the brake.

Therefore,

$\Sigma\mathfrak{I} = 0$ by Eq. (76), which (if moments are taken about the axis of the shaft) gives,

$$FL - fr = 0$$

or

$$fr = FL. \tag{b}$$

Substituting the value of fr from Eq. (b) in Eq. (a), we have:

$$P = 2\pi FLN \quad \text{gm-cm/min} \tag{163}$$

$$= \frac{2\pi FLN}{60} \quad \text{gm-cm/sec}$$

and at places where $g = 980$ cm/sec^2,

$$P = \frac{2\pi FLN\,(980)}{60} \quad \text{ergs/sec}$$

$$= \frac{2\pi FLN\,(980)}{60\,(10^7)} \quad \text{joules/sec or watts.} \tag{164}$$

Similarly, if N is in rpm; F is in lb force; and L is in ft:

$$P = 2\pi FLN \quad \text{ft-lbs/min}$$

$$= \frac{2\pi FLN}{33{,}000} \text{ hp.} \tag{165}$$

PROBLEMS

1. Sulphuric acid (sp. gr. 1.8) is used in an open manometer attached to a vessel. If the acid stands at 56.3 cm on the open side and 39.8 cm on the vessel side, when the barometer reads 75.1 cm of Hg, what is the absolute pressure in the vessel?

2. The difference in level of Hg on the two sides of an absolute gauge is 15 cm. What is the absolute pressure in the vessel to which the gauge is connected if the temperature is 19°C? What would an open manometer read at the same time if the barometer stands at 74 cm?

3. In a McLeod gauge a volume of 90 cm is compressed into a tube 1 mm in diameter and 8 mm long by a column of Hg 2 cm high. What was the original pressure in the bulb in mm of Hg and in baryes?

4. A diving bell 7 ft high inside is lowered until the top is 30 ft below the surface of the water. If the barometer reads 29.5 in. of Hg, how high does the water rise in the bell? What air pressure in the bell would be just sufficient to keep out the water?

5. A bubble rises to the surface from a depth of 30.6 meters. If the atmospheric pressure is 75 cm of Hg, what will be the relative volumes of the bubbles at the bottom and at the top?

6. A motor having a pulley 12 in. in diameter runs at 3000 rpm. What power is it developing when the spring balances of a strap brake on it read 55 lb and 25 lb, respectively?

7. An engine making 300 rpm is being tested with a Prony brake, the arm of which is 4 ft long. If the force on the scale is 143 lb, what power is the engine developing?

8. What power is developed by a motor making 1800 rpm if the Prony brake arm is 3 ft and the force on the scale is 132 lb?

9. The brake arm of a Prony brake is 5 ft long. If the wheel makes 250 rpm and the force on the scales is 150 lb, what power is developed by the engine?

CHAPTER XIII

ELASTICITY

171. Elasticity is defined as that property of matter by virtue of which it tends to return to its original shape or size after any change of shape or size. In general, any change in shape or size is called a *deformation.*

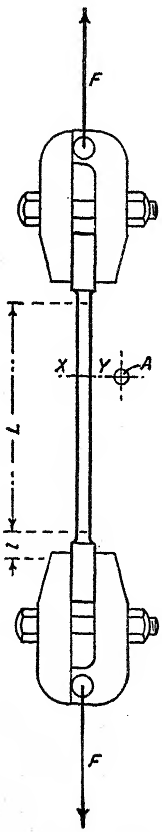


Fig. 150. Test Specimen

172. Strain is the ratio of the change in a quantity to the original quantity. It is often spoken of as the fractional, or percentage change, and is a pure number.

In the case of shear (see Sec. 178), strain is the tangent of the angular displacement of a line on the body.

Thus, when a body such as the test piece in Fig. 150 sustains a load F, its original length L is increased by an amount l which is the deformation—in this case an elongation. According to the definition, the strain is given by the equation:

$$\textbf{Strain} \equiv \frac{\text{Change of Length}}{\text{Original Length}}$$

$$= \frac{l}{L} \quad \text{(a pure number).} \qquad (166)$$

173. Stress. If a plane XY is passed perpendicular to the test piece at any point (Fig. 150), we may think of the part above the plane as holding up the part below and the load F by exerting an *internal force* upward at the plane. This is the force of cohesion.

Stress is defined as the internal force per unit of the area on which the force acts. It tends to resist the external load. For the

specimen under consideration,

$$\text{Stress} \equiv \frac{F}{A} \text{ dynes/cm}^2 \text{ or lb/in.}^2 \qquad (167)$$

according as cgs or B.E. units are used.

Engineers use the terms "unit strain" and "unit stress," respectively, instead of stress and strain as defined above.

174. Elastic limit. The maximum stress a substance will bear and yet return to its original shape or size when the load is removed is called its elastic limit.

If various loads are applied to the specimen and the corresponding elongations are recorded, it will be found that the specimen will return to its original length each time the load is removed, up to a certain value of the load; but for loads greater than this value, the specimen does not return to its original length when the load is removed.

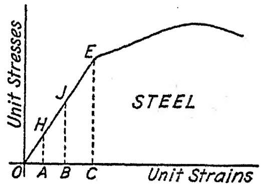


FIG. 151a. Stress-Strain Curve for Steel

If the strains and corresponding stresses are now computed and used as coordinates, a curve like that in Fig. 151a will be obtained. It will be found that ordinate CE of the point E, where the curve begins to deviate from a straight line, is the stress produced by the maximum load from which the specimen recovered its original length.

Therefore, EC represents the **elastic limit** of the specimen; i.e., the maximum stress after which the specimen returned to its original size when the load was removed.

175. Hooke's law. Since the right triangles of Fig. 151a are similar (having an acute angle in common), we have for any points H, J, on the straight part of the curve:

$$\frac{HA}{OA} = \frac{JB}{OB} = \frac{EC}{OC} \equiv k.$$

But HA, JB, and EC are stresses, and OA, OB, and OC are strains. Hence, the curve justifies Hooke's law: * **Up to the elastic limit, stresses are proportional to their corresponding strains.**

* Stated by Robert Hooke in 1676. Hooke suggested the use of hairsprings in watches (1665).

This is the fundamental law of elasticity. **A perfectly elastic body** would obey it **exactly.** Precise measurements * show that for actual substances it is only approximately true; but the approximation is sufficiently close for all practical purposes, such as the deflection of girders and the deformation of springs.

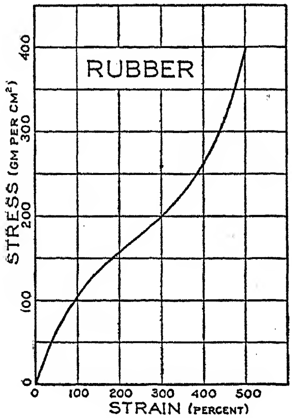


FIG. 151b. Stress-Strain Curve for Rubber. (Courtesy Dr. E. Guth, Univ. of Notre Dame, Ind.)

Figure 151b shows the stress-strain curve for rubber, which does not obey Hooke's law except for very small stresses.

176. Coefficient of elasticity. From Hooke's law, stresses y are proportional to their corresponding strains x up to the elastic limit; i.e.,

$$y \propto x$$

or

$$y = kx; \quad k = \frac{y}{x}.$$

The proportionality factor k is called a coefficient of elasticity. It is the slope of the line OE (Fig. 151a); i.e., $k = \tan\theta$. Or we may say: **A coefficient of elasticity is the ratio of a stress to its corresponding strain.**

There are only two coefficients of elasticity which are independent of each other, the volume coefficient and the coefficient of shear. All others depend upon these.†

177. Volume coefficient of elasticity, or bulk modulus, B is the ratio of the increase of pressure p to the volume strain v/V which it causes. Algebraically,

$$B \equiv \frac{p}{\left(\dfrac{v}{V}\right)} = \frac{pV}{v} \tag{168}$$

where V is the original volume and v is the change of volume due to the stress p.

* See article by J. O. Thompson, *Science* (Sept. 24, 1926).
† E. G. Coker and L. N. G. Filon, *Photoelasticity* (Cambridge, Eng., The University Press, 1931), p. 116.

To determine the volume coefficient of elasticity for a solid, the specimen S (Fig. 152) is placed in a very strong vessel V having heavy glass windows, through which its change of dimensions may be observed. The vessel is filled with a transparent liquid, and the hydrostatic pressure, applied by a pump, is indicated on the pressure gauge P.

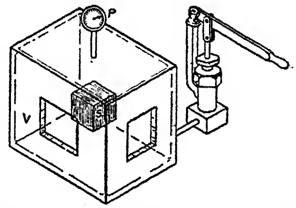


FIG. 152. Determination of Bulk Modulus

It will be noticed that volume coefficient, or bulk modulus, of elasticity gives the ratio of stress to strain in cases in which there is a **change of size without change of shape.**

178. Coefficient of shear, or modulus of rigidity. If the hand is placed with the palm flat upon a stack of cards or a thick book lying on a table, by pushing horizontally the rectangular pile may be changed into a parallelopiped (each card or page sliding over the one beneath). Here we have a change of shape without change of volume, since the altitude and the area of the base of the original and the final objects are the same.

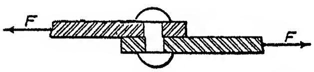


FIG. 153. Shear in a Riveted Joint

When an elastic body such as a block of jelly or steel is deformed in this way, its **shape being changed without change of volume,** it is said to undergo a **shearing strain,** or **shear,** and the internal forces between parallel layers tending to restore it to its original shape are called **shearing stresses.** The name comes from the fact that when a piece of pasteboard, say, is severed by a pair of shears (scissors), precisely this kind of stress is produced. Also, when a lap-riveted joint (Fig. 153) is stretched, the two plates tend to shear off the rivet, causing the upper half to slide off the lower half.

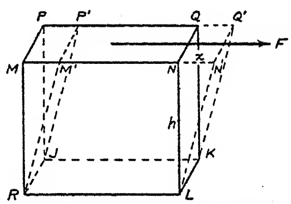


FIG. 154. Shearing Strain

Let the rectangular block of Fig. 154 have its base $JKLR$ fastened to the top of the table; and let a force F, uniformly distributed over surface $PQNM$, distort the block into the parallelopiped $P'Q'N'M' — JKLR$. Then each horizontal plane of the body is

made to slide somewhat over the next lower plane and the body experiences a **shearing strain,** i.e., a change of shape without change of volume.

To obtain an expression for the coefficient of elasticity for this kind of strain, we have:

$$\textbf{Shearing stress} \equiv \frac{F}{A}$$

where A is the area of the plane face $PQNM$ over which F is uniformly distributed. Also,

$$\textbf{Shearing strain} \equiv \frac{x}{h}.$$

So, from the definition of a coefficient of elasticity,

$$\textbf{Coefficient of shear,} \quad \eta \equiv \frac{\left(\frac{F}{A}\right)}{\left(\frac{x}{h}\right)} = \frac{Fh}{Ax} \tag{169}$$

which will be in dynes/cm^2 or in lb/in.^{2}, depending upon whether cgs or B.E. units are used.

179. Young's modulus. This coefficient of linear stretch is used so much in engineering tests that it is commonly referred to as "the coefficient of elasticity," or modulus of elasticity, E. It does not take into account the change of volume that occurs. A test piece arranged in clamps for the determination of its *Young's modulus** is shown in Fig. 150, where

$$\text{Tensile stress} \equiv \frac{F}{A}$$

$$\text{Tensile strain} \equiv \frac{l}{L}.$$

Hence, by the definition of a coefficient of elasticity,

$$\textbf{Young's modulus,} \quad E \equiv \frac{\left(\frac{F}{A}\right)}{\left(\frac{l}{L}\right)} = \frac{FL}{Al} \tag{170}$$

* After the English physicist Thomas Young (1773–1829).

and the value will be in dynes/cm² or in lb/in.² according as cgs or B.E. units are used.

COEFFICIENTS OF ELASTICITY

Substance	Young's Modulus Dyne/cm²	Young's Modulus Lb/in²	Modulus of Rigidity Dyne/cm²	Modulus of Rigidity Lb/in²	Bulk Modulus Dyne/cm²
Aluminum	7×10^{11}	10×10^6	2.5×10^{11}	3.6×10^6	7×10^{11}
Copper (hard)	10.	14.	4.2	6.1	12.
Iron (drawn)	20.	29.	8.	11.6	15.
Iron (cast)	11.5	16.8	5.1	7.4	9.6
Steel (mild)	22.	32.	8.	11.6	16.
Wood (mean)	.5	.75			
Water					.2
Mercury					2.6

Solved Problem

What will be the elongation of a steel test specimen, originally 1 ft long and 0.75 in. in diameter, when subjected to a tension of 5500 lb force?

From a table of physical constants, we find Young's modulus for ordinary mild steel to be 32×10^6 lb/in.²

Solving Eq. (170) for l,

$$l = \frac{FL}{AE}$$

where the value of $A = \frac{\pi}{4}\left(\frac{3}{4}\right)^2$.

Therefore,

$$l = \frac{5500 \times 12}{\frac{3}{4}\left(\frac{\pi}{4}\right)^2 \times 32 \times 10^6} = 0.0046 \text{ in.}$$

180. Beams. The great importance of Young's modulus in structural design may be seen by the following instance. It is shown in texts on the strength of materials * that the deflection d of a beam (Fig. 155) is given by the equation:

$$d = \frac{Wl^3}{4h^3bE} \qquad (171)$$

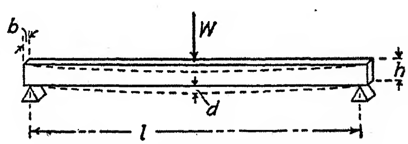


FIG. 155. Deflecting Beam

* Mansfield Merriman, *Mechanics of Materials* (New York, John Wiley & Sons, 1905), p. 75.

where W is the load;

l is the length;

h is the depth;

b is the breadth; and

E is Young's modulus,

for the material of which the beam is made.

Here it is seen that the **deflection varies inversely as the cube of the depth;** hence floor beams are invariably set on edge.

The elastic properties of the material of which the beam is made are represented in Eq. (171) by Young's modulus E. In specifications for structural materials, this modulus is stipulated more often than any other coefficient of elasticity.

PROBLEMS

1. If Young's modulus for iron is 30×10^6 lb/in², what elongation of a wire 10 ft long and 0.05 in. in diameter will be produced by a load of 50 lb?

2. The modulus of elasticity of a steel wire is 32×10^6 lb/in². If the wire is 8 ft long and 0.1 in. in diameter, what load will elongate it 0.25 in.?

3. What force will produce an elongation of 0.25 in. in a steel wire 30 ft long and 0.05 in. in diameter, if Young's modulus for steel is 30×10^6 lb/in.²?

4. A piece of brass wire 30.3 in. long and 0.0808 in. in diameter is elongated 0.0432 in. by a load of 100 lb. What is Young's modulus for this kind of brass?

5. If 2,000 in.³ of water are reduced to 1980 in.³ by a pressure of 3,000 lb/in.², what is the value of the bulk modulus for water?

6. Assuming the maximum depth of the sea to be 30,000 ft. and the mean density of sea water to be 1.03 gm/cm³, what would be the volume at this depth, of a piece of cast iron whose volume at the surface was 1 cm³? Would it float?

CHAPTER XIV

MOLECULAR FORCES

181. The inverse square law of gravitation, which, as we have seen in Sec. 122, predicts the mutual attraction of large (or *molar*) masses, does not account for the relatively great forces that act between molecules of matter at distances of the order of magnitude of 15×10^{-8} cm. This is called the approximate **range of molecular attraction.**

182. Cohesion and adhesion. The force that holds together molecules of the same kind is called *cohesion;* e.g., the force necessary to pull apart a piece of copper must overcome the cohesion between the molecules of copper.

In engineering, cohesion is called **tensile strength.**

Adhesion is the force that holds together molecules of different kinds. Thus, when a piece of pine is veneered with walnut, the two are united by a thin layer of glue. The glue adheres to the pine on one side and to the walnut on the other side; hence it is called an adhesive. If a glass disk *D* (Fig. 156) is lowered gently onto the surface of water in a dish, the water seems to grab it when it touches, and considerable force is necessary to pull the disk loose. When finally removed, it is still wet. The adhesion of water to glass is greater than the cohesion of water.

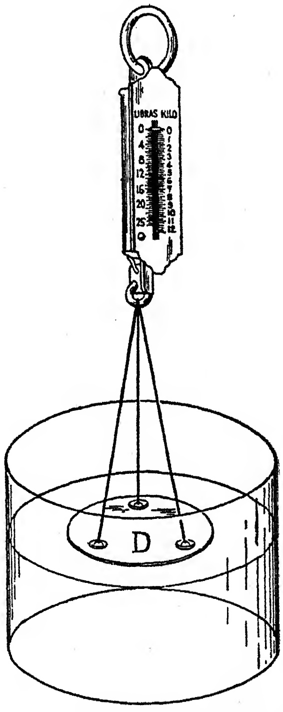


FIG. 156. Adhesion of Water to Glass

If mercury is used in the dish instead of water, more force is required to separate the surfaces than in the former case; but no

mercury sticks to the glass. The cohesion of mercury is greater than the adhesion of mercury to glass.

183. Absorption and adsorption. In this connection it is well to distinguish the terms "absorption" and "adsorption." **Absorption** is a case of true solution (see Sec. 198), the molecules of one substance intermingling with the molecules of another substance. **Adsorption,** on the other hand, is the adhering of the molecules of one substance to the surface of the other substance; it is a surface phenomenon.

For example, gold absorbs mercury and becomes very brittle; water absorbs air and other gases. But glass adsorbs air and water vapor. Charcoal, on account of its fine porous structure, adsorbs large amounts of all gases except neon and helium, and is therefore used in gas masks. It has been estimated that in consequence of the fine capillaries in coconut charcoal, 1 cm^3 has an exposed surface of 1/4 acre. It will adsorb about 82 cm^3 of ammonia at standard conditions. Platinum absorbs hydrogen and adsorbs oxygen. Both processes liberate heat and the hydrogen and oxygen are brought into intimate contact, so that spongy platinum is used in cigar lighters.

184. Surface tension. A molecule A (Fig. 157) in the interior of a liquid is attracted equally in all directions by its neighbors within the range of molecular attraction. A sphere described about A with this radius is called the **sphere of molecular attraction,** and the influence at A of molecules outside this sphere is negligible.

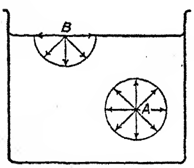


FIG. 157

But at a point B in a free surface against air or vapor, cohesive forces act on the molecule downward and sidewise only. The result of this inward attraction on the molecules at the surface is to crowd the liquid into the shape **having the least area possible under the circumstances.**

Hence raindrops, and drops of mercury on paper, are nearly spherical in form (the sphere having the least surface area for a given volume), and behave as if enclosed in a thin elastic membrane. Drops of olive oil, in a mixture of water and alcohol having the same density as the oil, are relieved from gravity and take a spherical form, standing anywhere in the vessel. Likewise, the

surface of a liquid in a dish behaves as if it were a stretched elastic sheet. And if a dry steel needle (sp. gr. 7.8) is carefully laid upon a surface of water (Fig. 158), it will be borne up in apparent defiance of Archimedes' law.

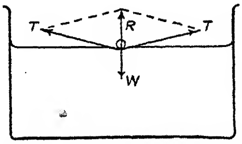


FIG. 158. Floating Needle

These and other related phenomena are due to the fact that the molecules in the free surface experience the force of cohesion downward and sidewise but not upward. They are most readily explained in terms of what is called surface tension.

Surface tension is the contractile force in the surface of a liquid by virtue of which the surface tends to shrink and assume the smallest area possible under the circumstances.

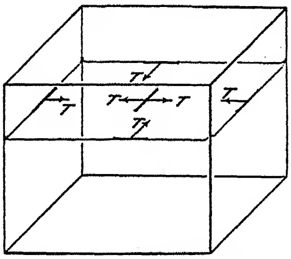


FIG. 159. Surface Tension

Surface tension T is measured in dynes per cm of length of the restraining edge, or of an imaginary line in the surface. Its direction is perpendicular to that edge or line (Fig. 159).

The floating of the needle may now be explained by saying that its weight W is supported by the resultant R of the equal forces T, T, on opposite sides of the depression which it makes in the surface.

In general, the magnitude of surface tension depends upon the nature of the two substances in contact at the surface.

185. Surface energy. The difference between surface tension and the tension in a stretched elastic sheet is made clearer by considerations of energy. For a particle C (Fig. 160) whose sphere of attraction is just tangent to the surface, cohesion is equal in all directions. But for one at D or at H the total force of cohesion downward is greater, say, than the total adhesion upward. To pull a molecule from C to H will require an upward force equal to the difference between these two;

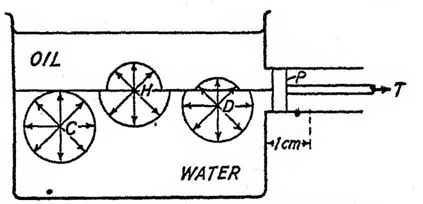


FIG. 160. Surface Energy

and since the molecule is moved through distance by this force, **work is done.**

Hence, work is required to draw molecules into a liquid surface, and therefore an amount of energy (approximately) equal to the work done is stored in the surface.

We have seen in Sec. 80 that a system always tends to adjust itself so that its potential energy is a minimum. We should therefore expect the surface to tend to contract, somewhat like a stretched elastic membrane, and pull the piston back. And so it does. But there is this great difference: As an elastic sheet contracts, its tension becomes less (see Hooke's law), and the contraction ceases when the sheet reaches a certain size—its unstretched area. The number of molecules (of rubber) in the sheet is always the same. But the surface of a liquid contracts because the cohesion of molecules below the surface pulls more and more of those in the surface down into the liquid as at C; the surface tension stays constant; and the contraction will continue indefinitely. A soap film across the large end of a funnel will contract and ascend to the top of the cone.

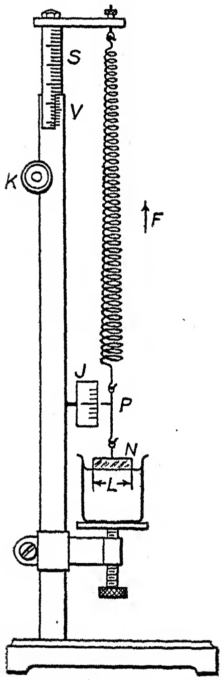


Fig. 161

If the frictionless piston P is 1 cm^2 (hydrostatic forces being balanced out), and a force equal to the surface tension T dynes pulls it 1 cm to the right, the liquid surface is increased in area by 1 cm^2, and the work done is T ergs.

Hence the energy of a liquid surface in ergs/cm^2 is numerically equal to its surface tension in dynes/cm.

186. Measurement of surface tension. A simple and fairly accurate method of measuring surface tension is shown in Fig. 161. A rectangular frame N of thin platinum wire is suspended from the spring of a Jolly balance, and the apparatus is adjusted until the horizontal top of the frame is a few millimeters above the surface when the pointer P stands opposite the mark fiducial on J.

The frame is then submerged in the liquid. On being withdrawn by turning the knurled head K, it is covered by a rectangular sheet of the liquid of the same length L as the frame.

This sheet of liquid has two surfaces which tend to contract and pull the frame downward. To overcome this surface tension, the knurled head K is again turned until the pointer is brought back to its original position. The elongation of the spring, read on the scale S by the vernier V, is a measure of the force F in dynes necessary to overcome the surface tension T of *two* surfaces, each of length L cm. Since T is defined as force per unit of length, the force due to each surface is TL dynes, and

$$F = 2TL$$

whence

$$T = \frac{F}{2L} \text{ dynes/cm.} \quad (172)$$

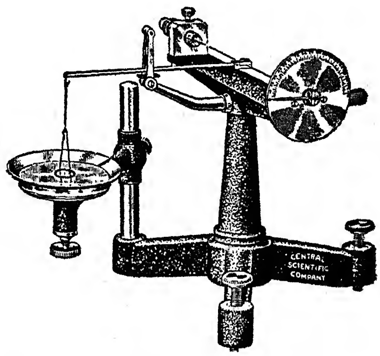


FIG. 162. DuNouy's Torsion Balance. (Courtesy Central Scientific Co.)

Much more accurate results may be obtained by the use of the Cenco-du Nouy torsion balance, shown in Fig. 162. The principle is the same as in the Jolly balance method, but a ring is used instead of a fork and the force is measured by twisting a wire, the angle of twist being proportional to the surface tension.

187. A table giving surface tensions follows.

TABLE OF SURFACE TENSIONS

Substance	In Contact with	Dynes/cm at 20°C
Alcohol, ethyl	Air	22.27
Benzene	Air	28.88
Glycerine	Air	63.14
Mercury	Air	476.0
Mercury	Water	375.0
Soap-water (strong solution)	Air	26.±
Turpentine	Air	27.2
Water	Air	72.75
Water	Benzene	35.0

188. Effect of temperature on surface tension. In general, surface tension decreases when temperature increases. Over moderate ranges the relation is linear, as is shown by the following equation:

$$T = T_0(1 - at) \tag{173}$$

where T is the surface tension in dynes/cm at temperature t°C;
T_0 is the surface tension in dynes/cm at 0°C; and
a is the temperature coefficient.

Figure 163 shows the surface tension-temperature curve for water.

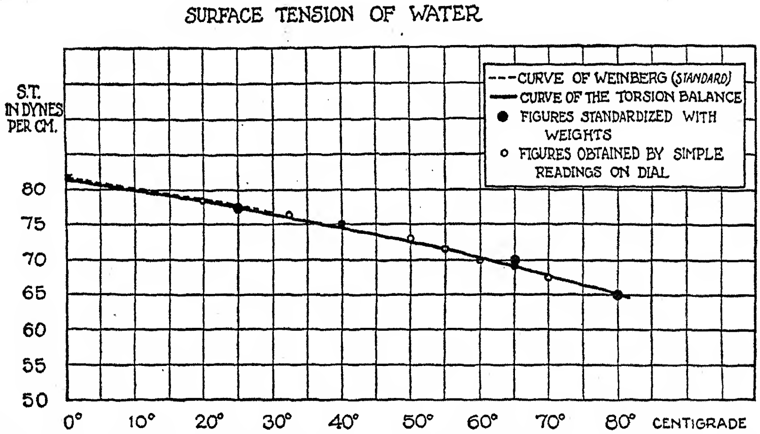


FIG. 163. Variation of Surface Tension with Temperature. (Courtesy Central Scientific Co.)

189. Some practical aspects of surface tension. The glass blower employs surface tension to round off the sharp edges of rods and tubes by merely heating them until they soften.

Soap added to spraying mixtures reduces their surface tensions and causes them to spread more readily over leaves and bark. One is rather surprised to find that the surface tension of a soap solution is less than that of pure water.

The spreading quality of paint and the value of solder depend upon their surface tensions. An antiseptic must penetrate dressings and crevices of wounds; hence its surface tension must be low. Type metal makes sharp outlines because of low surface tension when in the molten state.

Fabrics are made waterproof by chemical treatment that increases the surface tension of water in contact with them.

190. Capillarity. The behavior of liquids in tubes of small diameter—capillary tubes (from Latin, **capillus** = a hair)—is a consequence of surface tension. If adhesion between the liquid and the tube is greater than the cohesion of the liquid, the liquid wets the tube. If cohesion is greater than adhesion, the liquid does not wet the tube.

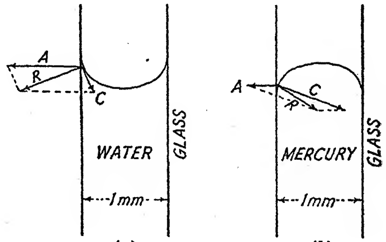


FIG. 164. Capillarity

In the case of water and glass (Fig. 164a), adhesion A is greater than cohesion C; whereas with mercury and glass (Fig. 164b), cohesion is greater than adhesion. In every case the liquid surface adjusts itself normal to the resultant force R. Hence the surface of water in glass is concave upward; and that of mercury, convex upward, as shown. The contraction of the surface tends to lift the water but to depress the mercury. The curved top surface is called the **meniscus.**

The phenomenon whereby liquids rise or are depressed in capillary tubes is known as capillarity, or capillary action. It is seen in the rise of oil in a lamp wick; and it partly accounts for the ascent of sap in plants. Water rises in the soil by capillarity. Cultivation in dry weather helps to conserve the moisture in the ground by breaking up the fine pores that would enable it to rise to the surface and evaporate.

ANGLES OF CONTACT

(International Critical Tables)

Liquid	Tube	Angle of Contact	Liquid	Tube	Angle of Contact
Alcohol	Glass	0°	Glycerine	Glass	0
Aqueous solution of various salts	Glass	0	Mercury	Glass	139
			Turpentine	Glass	0
			Water	Glass	0
Ether	Glass	0	Water	Paraffin	107

191. Angle of contact. The angle of contact θ between the surface of the capillary tube and the tangent to the liquid surface at the point of contact depends upon the kind of liquid and the material of the tube. The values obtained by different observers are in some cases widely at variance, no doubt because the surfaces of liquid and tube were not perfectly clean. The values in table on p. 217 are for clean surfaces, previously wet with the liquid.

192. Jurin's law. Consider a tube of radius r, not more than 1 mm, in which a liquid of density D is elevated a height h cm by capillary action. The lifting is due to the vertical component ($= T\cos\theta$) of the surface tension T, acting on each 1-cm length of the circumference of the circle where the liquid surface makes contact with the tube. Therefore,

$$\text{Total upward force} = 2\pi rT\cos\theta \quad \text{dynes.} \tag{a}$$

This upward force balances the weight W of the column of liquid elevated, which is:

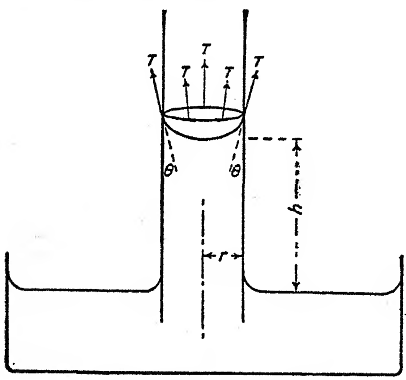


Fig. 165. Jurin's Law

$$W = \pi r^2hD \quad \text{gm force}$$
$$= \pi r^2hDg \quad \text{dynes.} \tag{b}$$

Equating the values from Eqs. (a) and (b),

$$2\pi rT\cos\theta = \pi r^2hDg$$

whence we have the equation for **Jurin's Law:**

$$h = \frac{2T\cos\theta}{rDg}. \tag{174}$$

For liquids that are depressed, θ, measured as indicated in Fig. 165, is greater than 90°; hence $\cos\theta$ will be negative and h will be negative, indicating depression instead of elevation.

193. Laws of capillary action.

1. Liquids rise in tubes of 3-mm diameter* or less if they wet the tubes, and the surface is concave upward.

* In larger tubes the rise or depression is negligible.

2. Liquids are depressed in tubes of 3-mm diameter or less if they do not wet the tubes, and the surface is convex upward.

3. The height of elevation (or depression) varies directly as the surface tension of the liquid and inversely as the radius of the tube and the density of the liquid.

PROBLEMS

1. A platinum fork, the horizontal part of which is 2 cm long, is just withdrawn from the surface of a soap solution by a force of 0.55 gm. If the fork weighs 0.5 gm, what is the surface tension of the solution?

2. An oil (specific gravity 0.9) rises 4 cm in a capillary tube 0.5 mm in diameter. If the angle of contact is 20°, what is the surface tension?

3. A certain oil whose angle of contact against glass is 24° and density 0.95 gm/cm^3 rises 2.5 cm in a glass capillary 1 mm in diameter. What is the surface tension of the oil?

4. If the ring in a Cenco-du Nouy apparatus is 4 cm in circumference and weighs 1.45 gm, and if a force of 2000 dynes is required to pull it from the liquid surface, what is the surface tension of the liquid?

5. In Cenco-du Nouy's apparatus the ring and its suspension weigh 1.32 gm. If the ring is 2 cm in diameter and a force of 2200 dynes is required to withdraw it from the surface of water, what is the surface tension of water?

CHAPTER XV

KINETIC THEORY OF MATTER

194. The atomic theory of the constitution of matter, as proposed in 1803 by the English chemist, John Dalton, is one of the cornerstones of physical science. The meaning of the word *atom*, however, has changed. As used by the ancient Greeks, it meant "an indivisible particle of matter"; but the discoveries of the electron (1897), the proton (1911), and the neutron (1932) have made us abandon that meaning.

Atoms are now believed to be small particles of matter made up of electrons, protons, and neutrons, in various arrangements.

A substance all of whose atoms are alike* is called **an element.** There are probably 92 elements as exhibited in the well-known periodic table. All but two of these have been found in nature.

195. Atoms combine into groups called **molecules.** For example, two atoms H_1 of hydrogen form a molecule H_2 of hydrogen; and two atoms of hydrogen and one of oxygen form a molecule of water H_2O. But some elements when subdivided into atoms do not have the same properties that they had before the division. Thus, atomic hydrogen H_1 does not have the same properties as ordinary hydrogen gas H_2, which we say consists of molecules.

A molecule is the smallest portion of matter which contains the same elements in the same proportions as does the same material when in larger pieces.

In the present chapter we are concerned only with molecules.

196. Evidence of molecular motion: diffusion. When a gas cock is open, the odor of the gas, if it is ordinary city gas, can soon be noticed in all parts of the room, even if there are no air currents. Still more striking is the following experiment: Place a watch

* Except for slight differences of mass which distinguish "isotopes" of the same element (Sec. 508).

glass containing a little bromine at the bottom of a beaker and cover the latter with a sheet of glass. The brown bromine gas will soon diffuse throughout the air in the beaker, notwithstanding the fact that the molecules of bromine have about five times the mass of the oxygen and nitrogen molecules which compose the overlying air, so that gravity tends to keep the bromine at the bottom.

With liquids, if some copper sulphate solution (sp. gr. = 1.25) is carefully introduced at the bottom of a test tube of water (sp. gr. = 1) by means of a thistle tube, the blue color of copper sulphate will gradually be detected throughout the entire liquid, showing that the denser solution has diffused upward throughout the less dense solution.

Diffusion is found even among solids. After sheets of gold and lead have been in intimate contact for a long time, gold may be detected throughout the lead. Manufacturers of high-grade resistance boxes find that soft solder (Pb 30%, Sn 70%) will in time diffuse into the wires that it unites, causing an appreciable change in their resistances.

Mercury quickly diffuses into copper, brass, lead, etc., generally causing them to become very brittle.

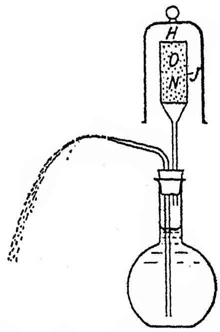


FIG. 166. Effusion of Gases

197. Effusion. If an inverted bell jar (Fig. 166) containing hydrogen is lowered over a porous jar J sealed to the top of a funnel whose tube is fitted into a water bottle, the hydrogen H_2 will pass through the pores and into the pot much more rapidly than the oxygen O_2 and the nitrogen N_2 will pass through the pores outward. The pressure on the surface of the water is consequently increased, and this forces out a lively stream from the nozzle.

198. Solutions. Many solids, when placed in various liquids, gradually disappear entirely as solids; but the presence of molecules of the substance of the solid may be detected by taste or by other tests in every part of the liquid—as when sugar is dissolved in tea.

A homogeneous molecular mixture is called a **solution**; the liquid is called a **solvent**; and the substance dissolved is called the **solute**.

If one continues to add the solute to the solvent, a state is reached after which no more of the solid will dissolve. The solution is then said to be **saturated.** If the temperature of the mixture is raised, more of the solid will, in general, go into solution; but eventually a new state of saturation will be reached at the new temperature. On cooling back to the original temperature, the additional amount of solid taken up will be precipitated, and the percentages of solvent and solute in the mixture will be the same as before heating.

199. Solutions of solids in solids, **solid solutions,** are familiar to the metallurgist. The composition of some very common ones is as follows:

Solid Solution	Composition	(percent)
Brass (commercial, yellow)....	67 Cu,	33 Zn
Bronze (gun metal)..........	90 Cu,	10 Sn
Coin Silver..................	90 Ag,	10 Cu
Coin Gold...................	90 Au,	10 Cu

200. Colloidal state. In true solutions as above described, the solute **breaks up into molecules** which distribute themselves in the interspaces between the molecules of the solvent. This is evidenced by the fact that the final volume of a solution, say, of alcohol and water, or of salt and water, does not equal the sum of the original volumes of the constituents. Hence liquids are porous: their molecules are not in contact.

A solid may, however, be finely divided and uniformly distributed throughout a fluid, without reaching the molecular subdivision of a true solution. Such a finely divided substance is said to be in the *colloidal state*, without specifying just how fine the division is.

Milk is such a colloidal suspension of butter fat in milk serum. Albumin, starch, glue, gum, and gelatin also give such suspensions. In fact, the name "colloid" comes from the Greek word for glue. But many substances not at all gluey may be put into the colloidal state. For example, gold, which has a crystalline structure, when very finely divided and distributed throughout glass gives the highest grade of ruby glass.

201. Semi-permeable membranes. The Scotch chemist, Thomas Graham, observed in 1861 that various animal and

vegetable membranes—such as the bladder, the inner skin of an egg, and the skin of a carrot, as well as membranes prepared by coating paper with a solution of starch or other colloid—would permit true solutions to diffuse through them readily but refused passage to substances in the colloidal state. Since crystalline substances generally give true solutions which pass the membranes, he called all substances that passed **crystalloids,** and all that were stopped, **colloids.** However, this distinction is no longer satisfactory, as, for example, in above case of gold.

Membranes which have the above property are called **semi-permeable membranes.**

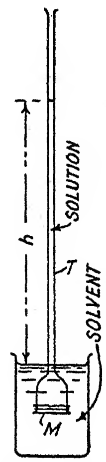


FIG. 167a. Osmosis

202. Osmosis. In 1854 Graham discovered that when two solutions of the same crystalloid but of different concentrations are separated by a semi-permeable membrane, the solvent passes from the weaker to the stronger solution while the crystalloid passes much more slowly from the stronger to the weaker solution, both processes tending to bring the two solutions to the same concentration. The phenomenon is known as *osmosis.*

Let the large end of a thistle tube T (Fig. 167a) be closed snugly by a piece of pig's bladder, or by a piece of filter paper that has been dipped in strong H_2SO_4; and let the tube be filled with a strong solution of copper sulphate or of cane sugar.

If the closed end is then submerged in a beaker of water, the water will diffuse through the membrane into the tube more rapidly than the solution will pass out. Hence the surface of the liquid in the tube will rise until the back pressure due to the column of solution h is sufficient to produce equilibrium; i.e., the state when as many molecules are passing in one direction as in the other.

The molecules of the crystalloid, passing through the membrane more slowly than the molecules of the pure solvent, gradually equalize the concentrations on the two sides of the membrane, while at the same time the height h decreases until the two solutions are again at the same level.

The pressure which causes the solvent to pass from the less

concentrated to the more concentrated solution is called the **osmotic pressure.** It is measured by the height h of the column of solution that is just sufficient to stop the action.

Artificial semi-permeable membranes have been prepared by H. N. Morse of the Johns Hopkins University, with which osmotic pressures as great as 400 lb/in.^{2} have been measured.

Osmosis cannot be satisfactorily explained on the "sieve theory"—i.e., on the assumption that molecules smaller than the pores in the membrane are permitted to pass, while the larger ones are stopped—as the following experiment shows. If wood alcohol and ether are separated by a rubber membrane, ether passes by osmosis into the alcohol. But if the same liquids are separated by a sheet of pig's bladder, osmosis takes place in the opposite direction; i.e., alcohol passes into the ether.

Osmosis occurs widely in nature. Plants take up moisture from the earth by osmosis, which with capillary action accounts at least to a considerable extent for the rise of sap. Water will pass by osmosis through the skin of a prune or a dried raisin until it assumes its original size; and "wieners," while cooking, will take up water until they burst.

In animals oxygen passes from the lungs into the blood, and carbon dioxide passes in the reverse direction, by osmosis through the membranes. Similarly, the products of digestion pass through the walls of the intestines by osmosis.

203. Laws of osmotic pressure. If we restrict ourselves to dilute solutions which are not ionized or in which chemical action does not take place, the following laws may be stated:

1. At a given temperature, osmotic pressure P is proportional to the concentration of the solution, i.e., to the number N of molecules of mass m per unit volume V.

$$P \propto \frac{Nm}{V}; \qquad PV = C'Nm = C. \tag{175}$$

2. For a given concentration M/V, osmotic pressure P is proportional to the absolute temperature T (see Sec. 263).

$$P \propto T; \qquad P = C''T. \tag{176}$$

3. At the same temperature, different solutions in which the masses of the dissolved substances are proportional to their

molecular masses exert equal osmotic pressures. For example, the molecular mass of grape sugar (glucose) is 180, and that of cane sugar (sucrose) is 342. When 180×10^{-3} gm of glucose and 342×10^{-3} gm of sucrose are dissolved in water so that each solution has a volume of 1 liter, the osmotic pressure of each solution is 22.4×10^{-3} atmosphere.

Since these solutions contain the same number of molecules per unit volume and at the same temperature produce the same osmotic pressure, the third law may be restated:

At the same temperature and osmotic pressure, equal volumes of different solutions contain equal numbers of molecules.

From the foregoing, it will be seen that of these laws:

The first is equivalent to Boyle's law (Sec. 165).

The second is equivalent to Charles' law (Sec. 264).

The third is equivalent to Avogadro's law (Sec. 207).

Hence in general, for dilute solutions which are not ionized, osmotic pressure obeys the gas laws.

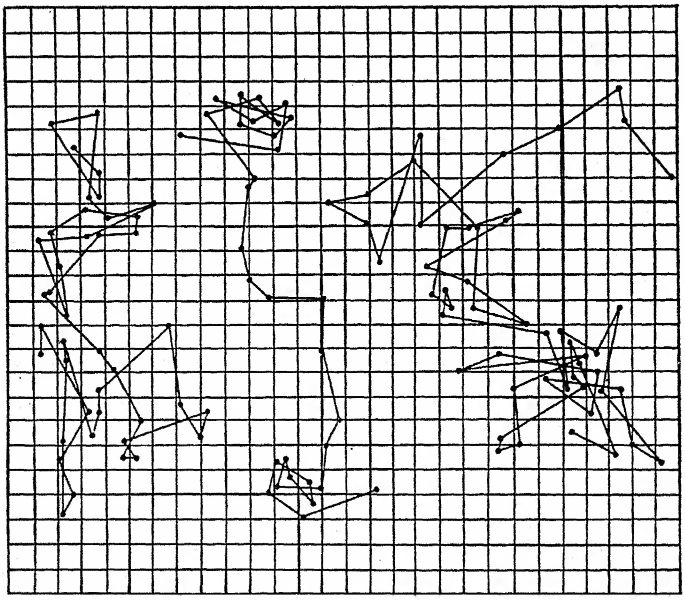


FIG. 167b. Graph of Brownian Movements. (After Perrin)

204. Brownian movements. Perhaps still more convincing evidence that molecules are in motion is afforded by the discovery

of Robert Brown, an English botanist, in 1827. He found that very small particles of foreign matter in suspension, when viewed with a microscope of moderate magnifying power,* can be seen to be continually in irregular motion, first in one direction and then in another, with various speeds (Fig. 167b).

This phenomenon is explained if we assume the molecules of the liquid to be moving about among themselves in all directions, colliding with and rebounding from one another in an entirely haphazard way. Then, at any given instant, a particle of the substance in suspension is bombarded on one side by, say, 10,000,000 molecules, while at the same instant 10,010,000 molecules strike it on the opposite side. It moves in the direction of the greater impact. A moment later it is driven in some other direction by the unbalanced impulses of two other swarms of molecules.

In 1905, Einstein succeeded in representing mathematically the Brownian motion of particles, and his results have been completely verified by experiment.

205. The kinetic theory of matter. The evidence of diffusion, effusion, osmosis, and Brownian movements leads us to believe that the molecules of all matter are in motion. This doctrine, known as the kinetic theory of matter, receives its support from the experimental confirmation of its predictions.

Mean free path. The average distance which a molecule travels between successive collisions with its neighbors is called its **mean free path.**

In solids, the mean free path is very short, the molecules resembling on a minute scale people in a very densely packed crowd. Each individual can move slightly, but his motion is restricted by contact with his fellows to the immediate neighborhood of a mean position.

In liquids, the mean free path is longer than in solids; and, with many collisions, molecules may wander through the entire mass like dancers in a crowded ballroom.

In gases, a molecule's mean free path is relatively very large (about 0.00001 cm or several hundred molecular diameters), and the molecules dart about in all directions with very great speed.

* Jean Perrin, *Atoms* (London, Constable & Co., 1923), p. 83.

For example, the speed of hydrogen molecules at standard conditions is 1839 m/sec.

The speeds of gas molecules under ordinary conditions are of the order of magnitude of the speed of rifle bullets. It would be possible to sustain a plate of steel in the air by firing bullets from many machine guns against its lower surface. In a similar way, the pressure of a gas against the walls of its containing vessel is believed to be due to **molecular bombardment.**

When two molecules of a gas approach each other very closely, their motions are opposed by repulsive forces—probably because of their electrical nature—which increase until they are sufficient to prevent further approach. The molecules then rebound and move in straight lines until they again collide with other molecules or with the walls of the containing vessel.

Assuming that, in these collisions, the energy lost by one molecule is gained by another, and that the molecules rebound from the walls with the same speed with which they strike, the average linear kinetic energy ($1/2\ mv^2$) is constant and hence the average v^2 remains constant.

Daniel Bernoulli (1700–1782) demonstrated that, on these assumptions, the kinetic theory leads to Boyle's law (Sec. 165). It may be shown in an elementary way as follows:

Consider a mass D of gas confined in a cubical vessel (Fig. 168), 1 cm on each edge. Let m be the mass of each molecule, and n the number of molecules in the cubic centimeter. The density is then

$$D = nm. \qquad \text{(a)}$$

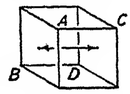


FIG. 168

There are three axes perpendicular respectively to the pairs of opposite faces of the cube, and the probability is as great that a molecule will be moving parallel to one of these axes as to another. Hence we may assume that $n/3$ molecules are moving perpendicular to each face, or the equivalent of this, at any instant.

Let v cm/sec be the velocity of a certain molecule perpendicular to face AB. Then the time for this molecule to make the round trip between two impacts against AB will be:

$$\frac{2\ \text{cm}}{v\ \text{cm/sec}} = \frac{2}{v}\ \text{sec}$$

and the **number of impacts** per second of this molecule against AB will be:

$$\frac{1 \text{ sec}}{\frac{2}{v} \text{ sec}} = \frac{v}{2}. \tag{b}$$

At each impact, the momentum of the molecule changes from $+mv$ to $-mv$, the total change being therefore $2\,mv$. And since there are $v/2$ impacts per second, the *total change of momentum* of the molecule in 1 sec is:

$$\frac{v}{2}(2mv) = mv^2. \tag{c}$$

Since this change of momentum takes place in 1 sec, it is the *time rate of change of momentum* of the molecule, and is therefore (Sec. 37) the average force exerted by one molecule on the face AB.

If we call the **average squared velocity** $\equiv \bar{v}^2$, then the **total force** exerted by the gas on the face AB is:

$$F = \frac{n}{3} m\bar{v}^2 \text{ dynes}. \tag{d}$$

And since the area is 1 cm², it is also the *pressure* (P).

Therefore

$$P = \frac{nm}{3}\bar{v}^2 \text{ dynes/cm}^2 \tag{177}$$

and by Eq. (a),

$$= \frac{D}{3}\bar{v}^2 \text{ dynes/cm}^2. \tag{178}$$

But by definition, $D \equiv M/V$, where M is the total mass and V the total volume of any body. Hence, for any body of gas whatever,

$$P = \frac{1}{3}\frac{M}{V}\bar{v}^2$$

whence

$$PV = \frac{1}{3}Nm\bar{v}^2 \tag{179}$$

where N is the total number of molecules in the volume V.

In Sec. 269 it will be shown that if the temperature is constant, the average kinetic energy of a molecule ($1/2\ m\bar{v}^2$) is constant also.

Hence Eq. (179) may be written:

$$PV = \text{Constant, provided } T \text{ is constant,}$$

which is Boyle's law.

Solved Problem

If the density of hydrogen gas at S.T.P. (standard temperature and pressure) is 0.00008988 gm/cm³, what is the ("effective") velocity of the molecules?

Known:

$$\begin{aligned}
\text{Density of } H_2 &= 0.00008988 \text{ gm/cm}^3 \\
1 \text{ std. atmos.} &= 76 \text{ cm of Hg} \\
&= 76 \text{ cm} \times 13.6 \frac{\text{gm}}{\text{cm}^3} = 1033 \frac{\text{gm force}}{\text{cm}^2} \\
&= 1033 \frac{\cancel{\text{gm fce}}}{\text{cm}^2} \times 980.7 \frac{\text{dyne}}{\cancel{\text{gm fce}}} \\
&= 1{,}013{,}000 \text{ dyne/cm}^2\text{, or baryes.}
\end{aligned}$$

From Eq. (178), $\qquad P = \frac{D}{3}\bar{v}^2.$

Solution: Substituting the above values in Eq. (178),

$$\begin{aligned}
1{,}013{,}000 &= \tfrac{1}{3}(0.00008988)\bar{v}^2 \\
\bar{v} &= 1839 \text{ m/sec.}
\end{aligned}$$

206. Maxwell's law. From the principles of mechanics and the theory of probability, James Clerk Maxwell, in 1859, showed by mathematical analysis beyond the scope of this text that the kinetic theory of matter leads logically to the conclusion which is now called Maxwell's law:

At the same temperature, the molecules of all gases have the same average kinetic energy except near the temperature of absolute zero.

Let m_1, m_2 be the masses, and $\bar{v}_1{}^2$, $\bar{v}_2{}^2$, the average squared velocities, respectively, of the molecules of any two gases. Then, in algebraic form, Maxwell's law is:

$$\tfrac{1}{2}m_1\bar{v}_1{}^2 = \tfrac{1}{2}m_2\bar{v}_2{}^2. \tag{180}$$

207. Avogadro's law. We are now in a position to deduce rationally the fundamental law which originated in 1811 as an inspiration of Amadeo Avogadro, professor of physics at Turin.

From Eq. (179),

$$P_1V_1 = \tfrac{1}{3}N_1m_1\bar{v}_1^2 \qquad \text{for one gas;}$$

$$P_2V_2 = \tfrac{1}{3}N_2m_2\bar{v}_2^2 \qquad \text{for another gas.}$$

If we adjust $P_1 = P_2$, and take $V_1 = V_2$, we have:

$$\tfrac{1}{3}N_1m_1\bar{v}_1^2 = \tfrac{1}{3}N_2m_2\bar{v}_2^2$$

and multiplying through by 3/2,

$$\tfrac{1}{2}N_1m_1\bar{v}_1^2 = \tfrac{1}{2}N_2m_2\bar{v}_2^2.$$

But by Maxwell's law, if the temperature is the same for the two gases,

$$\tfrac{1}{2}m_1\bar{v}_1^2 = \tfrac{1}{2}m_2\bar{v}_2^2.$$

Therefore,

$$N_1 = N_2$$

which is *Avogadro's law:* **Equal volumes of different gases at the same temperature and pressure contain the same number of molecules.**

This law was the basis for the original determinations of relative molecular and atomic masses as employed in chemistry. Today they are verified by means of the mass spectrograph.

208. Avogadro's number. Up to 1908 there still remained authorities who doubted the kinetic theory of matter. In that year, Jean Perrin,* from measurements made on Brownian movements, obtained a value for **the number of molecules in a gram molecule** of a substance, which agreed so closely with the predictions of kinetic theory as to convince the most skeptical. This number is called *Avogadro's number* (N_0). The accepted value, obtained by Millikan, is now:

$$N_0 = 6.064 \pm 0.006 \times 10^{23} \frac{\text{molecules}}{\text{gram molecule}}.$$

The computation of this number is given in Sec. 429.

209. Loschmidt's number. The number of molecules in one cubic centimeter of a gas at standard conditions of temperature

* Jean Perrin, op. cit., p. 105.

and pressure is called *Loschmidt's number* (n_0), after Joseph Loschmidt, a schoolmaster of Vienna, who first computed it in 1865. It is found by dividing Avogadro's number by the number of cubic centimeters in 1 gm-molecular-volume, i.e., the volume of 1 gm molecule at S.T.P. This volume is readily obtained by dividing the mass of 1 gm molecule by the density of the gas. Thus, for hydrogen,

$$\text{1 gm molecule} = 2.0156 \text{ gm}$$
$$\text{Density} = 0.00008988 \text{ gm/cm}^3.$$

Therefore,

$$\text{1 gm-molecular-volume} = \frac{2.0156 \text{ gm}}{0.00008988 \text{ gm/cm}^3} = 22{,}420 \text{ cm}^3.$$

Hence *Loschmidt's number* is:

$$n_0 = \frac{6.064 \times 10^{23}}{22{,}420} = 2.705 \times 10^{19} \frac{\text{molecules}}{\text{cm}^3}.$$

210. Size of molecules. Mean free path. By means of Loschmidt's number, we can easily compute the maximum size that a molecule of a gas can have, provided the molecules are all alike.

Since liquids require great force to compress them but slightly, we may assume that their molecules are practically in contact.

A cubic centimeter of water makes 1700 cm³ of steam at 100°C and 76 cm of Hg pressure. This volume reduces to 1240 cm³ at standard conditions (0°C and 76 cm of Hg pressure). Therefore,

Number of molecules in 1 cm³ of water

$$= 1240 \times (2.7 \times 10^{19}) = 33 \times 10^{21}.$$

Number of molecules along each edge of cm³

$$= \sqrt[3]{33 \times 10^{21}} = 3.2 \times 10^7.$$

Hence, if the molecules are all alike,

Maximum diameter of water molecule

$$= \frac{1 \text{ cm}}{3.2 \times 10^7} = 3 \times 10^{-8} \text{ cm} \equiv 3 \text{ Å}.$$

It may be shown similarly that the maximum diameter which a molecule of carbon dioxide can have is 4 Å; and that of a molecule of hydrogen, 3.6 Å; but these values should be thought of only as rough estimates.

If the molecules retain their size when water changes into steam, the unoccupied volume in 1240 cm³ would be 1239 cm³, so that the average distance between molecules in the gas state is large compared to the size of the molecules themselves.

The **mean free path** has been defined as the average distance a molecule travels between successive collisions with other molecules. Maxwell has computed* that under standard conditions, the mean free path of a molecule of a gas is about 0.00001 cm (= 1000 Å). Even with this great freedom of motion, a molecule of a gas under standard conditions is estimated to make 8×10^9 impacts per second upon its neighbors.†

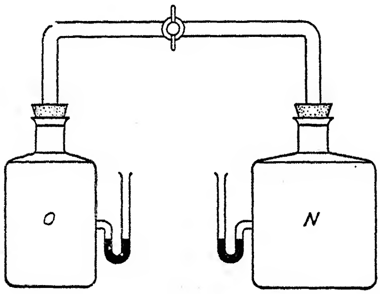


FIG. 169. Dalton's Law

211. Dalton's law of partial pressures. Consider a vessel O (Fig. 169), containing 1 liter of oxygen and connected by a gas-tight tube and valve to another vessel N, containing 4 liters of nitrogen, the valve being closed. Let both gases be at the same temperature and at a pressure of 1 atmosphere, as indicated by the manometers.

When the valve is opened, each gas will diffuse into the other vessel, being merely retarded by the gas already present, until the mixture becomes homogeneous. Each gas will then be distributed uniformly throughout the whole volume of both bottles; and the space between molecules is so great, compared with the size of the molecules themselves, that the molecules of each gas will bombard the container just as if the other gas were not present. The pressure in both containers will remain 1 atmosphere.

By Boyle's law we can calculate the pressure exerted by each gas after the expansion:

For oxygen,

$$P_o''V_o'' = P_o'V_o'; \quad V_o'' = 5V_o'; \quad P_o' = 1 \text{ atm.}$$

$$P_o'' = \frac{V_o'}{V_o''}P_o' = \frac{1}{5} \text{ atm.}$$

* Clerk Maxwell, *Phil. Mag.*, xxviii (1860).

† P. G. Tait, *Properties of Matter* (New York, The Macmillan Company), p. 25.

For nitrogen,

$$P_N''V_N'' = P_N'V_N'; \quad V_N'' = \frac{5}{4}V_N'; \quad P_N' = 1 \text{ atm.}$$

$$P_N'' = \frac{V_N'}{V_N''}P_N' = \frac{4}{5} \text{ atm.}$$

It will be observed that the sum of the two partial pressures equals the total pressure. A similar result will be obtained if we introduce into a vessel any number of gases that do not react chemically.

These facts are embodied in the **law of partial pressures** formulated by John Dalton, the founder of the atomic theory of chemistry, in 1802: **In a mixture of gases at a given temperature (provided the gases do not react chemically), each gas exerts the same pressure as if it alone occupied the entire volume at that temperature; and the total pressure is the sum of these partial pressures.**

It is also of interest that when a mixture of gases is in contact with the surface of a liquid, each gas will dissolve in the liquid according to its own partial pressure, quite regardless of the presence and the pressure of the other gases.

CHAPTER XVI

WAVE MOTION

212. Waves. Everyone is familiar with the waves that spread out in ever widening circles when a stone is dropped into a pond, and with the waves of the sea, whose enormous energy is shown by their destruction of ships and of shore lines. If one end of a heavy cord is fastened to a wall and the other end is moved up and down with a regular motion, a wave is seen to traverse the cord. In this case the distinctive feature is the temporary change of shape that moves through the medium, away from the source.

Quite as real, if not so realistic, are the radio waves—electromagnetic waves—by means of which the energy of the transmitter is radiated to receiving stations everywhere. These waves, like those of light and sound, are not visible, but they can be shown to consist of temporary changes of the condition of the medium, in which they move outward from the source.

A **wave** is a temporary change in the shape or condition of a medium, which moves through the medium and transmits energy outward from the source.

Waves generally follow one after another at equal intervals. Such a procession is called a **train of waves,** or a **system of waves.** A noteworthy instance of a single wave was the Lisbon earthquake wave of 1755, often erroneously called a "tidal wave," which took the lives of 30,000 persons.

213. Distinction between the motion of the wave form and of the particles of the medium. Let a piece of cork or a chip of wood be placed on the surface of a smooth body of water; and let a train of waves be started and maintained by moving a stick up and down rhythmically in the water, the surface of which is otherwise undisturbed. Any one wave will move outward in a circle of increasing radius. But the piece of cork, if observed closely, will be seen to rise with the hump of water, advance a short distance,

descend into the succeeding hollow, and retreat the distance that it advanced; and when the wave motion ceases, the cork will be just where it was at first. The float obviously partakes of the motion of the particles of the water at the surface; and careful experimentation reveals that in such water waves each particle describes a circle or an ellipse about a mean position. (See Fig. 172.) **Only the wave form proceeds continuously forward.**

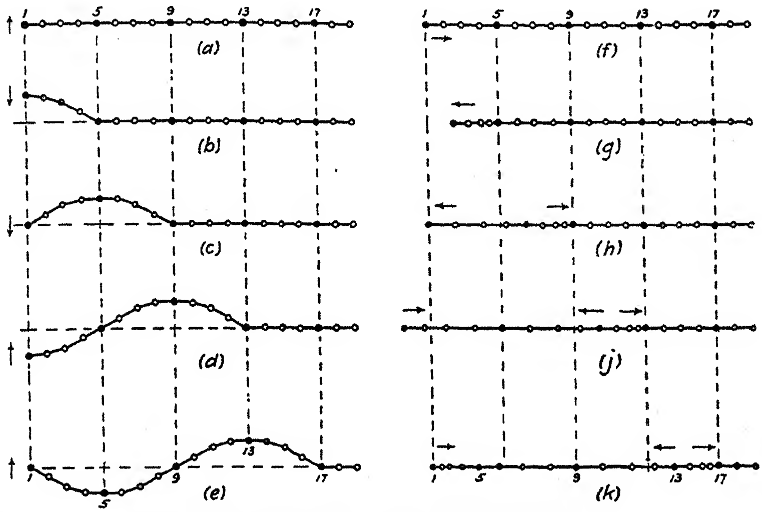


FIG. 170. Transverse and Longitudinal Waves

214. Kinds of waves. Waves are classified on the basis of how the particles of the medium move with reference to the direction of propagation of the wave form.

(*a*) *Transverse waves.* Let a number of buttons be attached to a long elastic cord at equal intervals, one end of the cord being attached to a rigid support, as in Fig. 170a. If button number 1 is then given an oscillatory motion at right angles to the cord, button number 2 will follow it due to the elastic force in the cord, but will lag somewhat behind it on account of inertia. Similarly, number 3 will lag somewhat behind 2, 4 behind 3, and so on. When number 1 has reached the highest point in its path, 5 will be just starting upward (Fig. 170b).

Number 1 then begins to move downward; but 2, on account of its inertia, continues upward until it rises to the height that 1 reached. It then begins to descend, while 3 continues upward, as

do 4 and 5; and 6 starts as 2 reaches the top. When 1 gets back to its initial position, moving downward, the cord has the form of Fig. 170c.

Number 1 continues downward, 2, 3, etc., following in order; and when 1 reaches a point as far below its initial position as it previously was above that position, the cord has the form of Fig. 170d.

Number 1 then starts upward again, and when it arrives at its original position (i.e., its position of equilibrium), the cord will have the form shown in Fig. 170e. Button number 1 is now back in its initial position, and is starting on a second vibration exactly like the first.

From this we see that **when a particle of the cord executes a complete vibration, a complete wave (one hump and one hollow) is set up in the cord.**

If the motion of 1 is maintained, all the particles of the cord (i.e., the medium) will acquire a similar motion; and a train of waves will move to the right in the cord. Such waves are called **transverse.**

A transverse wave is a wave in which the particles of the medium vibrate, or the condition of the medium changes, at right angles to the direction of propagation of the wave. We shall see presently that light and radio waves are transverse waves of the latter kind.

(*b*) *Longitudinal, or compressional, waves.* Instead of having button 1 vibrate at right angles to the cord, we may cause it to vibrate back and forth along the direction of the cord. Starting to the right, the portion of the cord between 1 and 2 will be compressed, and hence the motion will be communicated to 2, then to 3, and so on. When 1 reaches its extreme position to the right, the cord will have the appearance of Fig. 170g.

Number 1 will then have its motion reversed; but 2, 3, etc., will continue moving to the right for a time on account of inertia, and the motion of each will reverse when it has traveled as far to the right of its initial position as did number 1. Number 1 is now moving to the left, and since 2 lags behind it on account of inertia, the cord between 1 and 2 will be in tension and the buttons will begin to separate.

When 1 has reached its extreme limit of travel to the left, this

separation (rarefaction) will be a maximum (Fig. 170j). Number 1 then reverses its motion again, reducing the tension in the section of the cord between 1 and 2, until it returns to its original position and is ready to repeat its vibration (Fig. 170k).

It will be seen that during one complete vibration of a particle (button) of the cord, the particles are compressed, rarefied, and again compressed. If the motion is continued, these compressions and rarefactions will proceed along the cord and constitute a train of longitudinal waves, each wave consisting of one compression and one rarefaction.

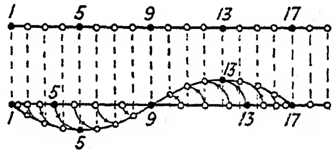


Fig. 171

A longitudinal, or compréssional, wave is a wave in which the particles of the medium vibrate, or the condition of the medium changes, in the direction of propagation of the wave. Sound waves will be seen to be an example of this type of wave.

Longitudinal waves are most conveniently represented to the eye by arbitrarily erecting at the equilibrium position of each particle a perpendicular equal to its displacement: upward if its displacement is positive, and downward if negative. When this convention is used, the curve representing a longitudinal wave is similar to that for a transverse wave; and the same discussion will apply to either kind of wave unless otherwise stated. The compressional wave of Fig. 170k has been treated in this way in Fig. 171, and the resulting graphical representation is seen to be identical with Fig. 170e.

(*c*) *Gravitational, or water, waves.* The waves ordinarily seen agitating the surface of a liquid are due to the weight of the liquid, and are therefore called gravitational waves, or water waves. A stone dropped into water makes a depression where it strikes and piles the water up slightly around the depression. The weight (gravity) of the surrounding water forces the depressed particles back to the original level; but reaching there with a velocity, their inertia carries them above that surface, thus producing an elevation where the depression was first made.

As was stated in Sec. 213, careful experiment shows that **in a gravitational, or water, wave the particles describe circles (or**

ellipses when the depth is small) about a mean position as a center. (See Fig. 172.)

At their highest and lowest positions, the particles are moving parallel to the direction of propagation; whereas at their extreme horizontal displacements, they are moving at right angles to that direction. Hence, water waves are a combination of longitudinal and transverse waves.

The part of the wave above the surface of equilibrium is called the **crest,** and the part below that surface, the **trough.**

Two other types of wave may occur in liquids:

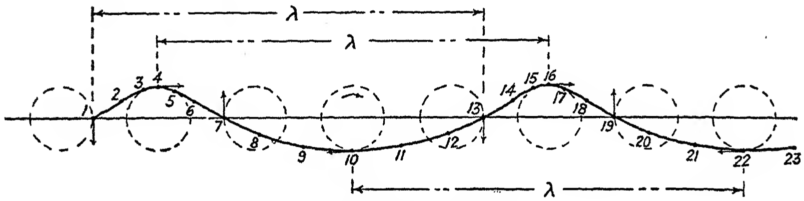


Fig. 172. Water Waves.

Ripples are very short waves (about 2 cm or less) for which the greater part of the restoring force is due to surface tension.

Compressional, or longitudinal, waves are transmitted by a liquid at a considerable distance below the surface, just as by any other ponderable elastic medium, but they are not visible like surface waves. Such waves are employed for depth soundings at sea.

215. Wave characteristics. The **amplitude** r of a wave is the maximum distance that a particle of the medium vibrates to either side of its equilibrium position.

The **period** T is the time required for a particle of the medium to make one complete vibration.

The **frequency** n is the number of complete vibrations which a particle of the medium makes in 1 sec. It is obviously the reciprocal of the period.

$$n = \frac{1}{T}\cdot \tag{181}$$

As was seen in the preceding section, while a particle executes a complete vibration, a complete wave is sent forward in the me-

dium. Hence, the **frequency** is also the number of waves that are produced in 1 sec, or the number of waves that pass a fixed point in 1 sec.

The **phase** of a particle participating in a wave motion is the fraction of a period that has elapsed since the particle last passed through its position of equilibrium moving in the positive direction. Phase may be expressed in time or in angle, which is proportional to time. The angle θ is that described by the radius OQ of Fig. 37, which turns about O with constant angular velocity ω and makes a complete revolution while a particle P of the medium makes a complete vibration. For example, in Figs. 170e and 170k the phases of several particles are as follows:

Particle No.	Phase in Time	Phase Angle
1	$1T$ or 0	360° or 0°
5	$3/4T$	270
9	$1/2T$	180
13	$1/4T$	90
17	0 or $1T$	0 or 360

The **wave length** λ is the distance from any particle to the next particle having the same phase. Thus, in Fig. 172, the wave length is from 1 to 13, or from 4 to 16, or from 10 to 22, etc.

By following the motion of particle 1 in Figs. 170e and 170k through a complete cycle, it will be seen that the **wave length is also the distance that the wave form advances while a particle makes one complete vibration, i.e., during one period.**

In transverse waves a wave length includes one crest and one trough; in longitudinal waves, one compression and one rarefaction.

The **wave form** is the curve obtained by drawing a line through the instantaneous positions of all the particles from one particle

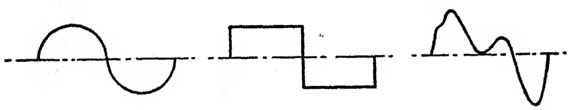


FIG. 173. Wave Forms

of zero phase to the next particle of zero phase. A wave can have almost any form one chooses (Fig. 173), but it may be difficult to design a mechanism to produce that form.

The simplest way to produce waves is from vibrating elastic bodies. Since the particles of an elastic body, vibrating freely, execute simple harmonic motion (very nearly), the motion of the particles in waves produced by vibrating elastic bodies will be practically shm. This is why waves are usually represented by simple periodic, harmonic, or sine curves. Such a curve is often called a "sine wave," but it is very difficult to secure a pure sine wave on account of the persistence of what are called "harmonics" (see Sec. 240).

The usual method of constructing a sine wave form is shown in Fig. 174. An auxiliary circle, having a radius r equal to the given

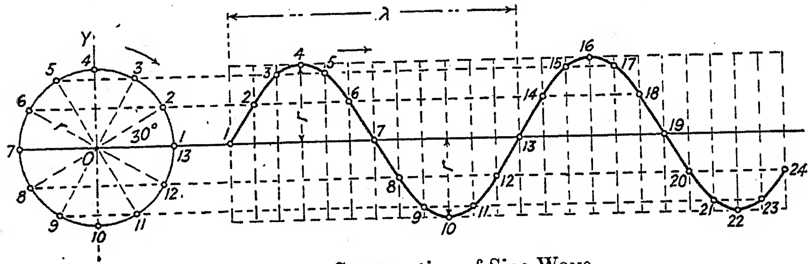


FIG. 174. Construction of Sine Wave

amplitude, is drawn with its center O on the center line of the wave; and its circumference is divided into equal arcs by as many points as are desired in the proposed wave, say, 12.

If the circle is imagined to rotate with the frequency which the wave is to have, the projections of these points on the vertical diameter OY will execute shm's successively differing in phase by 30°, which motions are the same as those of corresponding points of the wave.

Consequently, the given wave length λ is divided into the same number (12) of equal parts; and perpendiculars, erected at the points of division, will be the paths of vibration of the particles participating in the wave motion. Each point of the circle is then projected over to its corresponding path, and a smooth curve drawn through the points thus obtained is the required sine wave, or harmonic curve.

216. Velocity of waves. As was seen in Sec. 214, any wave form advances a distance equal to the wave length during one complete vibration of a particle of the medium. Hence the total dis-

tance the wave will advance in one second, i.e., the velocity v of the wave, is the product of the frequency n and the wave length λ:

$$v = n\lambda = \frac{\lambda}{T}. \tag{182}$$

This relation is true for any kind of wave.

Other expressions for the wave velocity may be obtained in terms of the kind of stress involved and the properties of the medium:

(a) Transverse waves in an elastic string or wire may be shown to have the velocity:

$$v = \sqrt{\frac{F}{M}} \tag{183}$$

where v is the velocity in cm/sec;
F is the tension in the string in dynes; and
M is the mass of unit length of the string in gm/cm.

(b) Longitudinal, or compressional, waves were shown by Newton to have the velocity:

$$v = \sqrt{\frac{E}{D}} \tag{184}$$

where v is the velocity in cm/sec;
E is the volume coefficient of elasticity for fluids, or Young's modulus for solids, in dynes/cm^2; and
D is the density of the medium in gm/cm^3.

(c) Waves on the surface of a liquid may be shown to have a velocity given by the following relation: *

$$v^2 = \frac{g\lambda}{2\pi} + \frac{2\pi T}{D\lambda} \tag{185}$$

where v is the velocity in cm/sec;
λ is the wave length in cm;
g is the acceleration due to gravity in cm/sec^2;
T is the surface tension in dynes/cm; and
D is the density of the liquid in gm/cm^3.

* W. Watson, *A Text-Book of Physics* (New York, Longmans, Green & Company, 1919), pp. 340, 354.

For wave lengths greater than 10 cm, the force of restitution is principally gravity, the surface tension term being negligible. Hence, for **gravitational waves,**

$$\left.\begin{aligned} v &= \sqrt{\frac{g\lambda}{2\pi}} && \text{in deep water} \\ v &= \sqrt{dg} && \text{in shallow water} \end{aligned}\right\} \qquad (186)$$

where d is the depth of the water.

But for waves shorter than 0.3 cm, surface tension predominates and the first term of Eq. (185) may be neglected, so that for these very short **capillary waves,**

$$v = \sqrt{\frac{2\pi T}{D\lambda}}. \qquad (187)$$

From Eq. (185) it will be seen that in water, long waves travel faster than shorter ones. Thus, waves 100 ft long travel 15.6 mph, and those 1000 ft long travel 49.5 mph. Longer waves are therefore constantly overtaking shorter ones, which, with the many different directions of propagation, causes the very complex wave systems seen at sea.

217. Huyghens' principle. When a wave is propagated in any medium, there are at any instant certain points that are just beginning to take part in the motion. The **wave front** is defined as the surface which contains all these points that are just being reached by the disturbance at that instant.

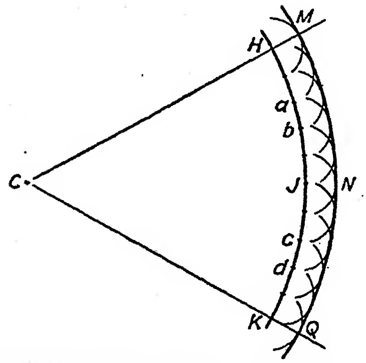


FIG. 175. Huyghen's Principle

Huyghens' principle states that every point on a wave front may be considered a new center of disturbance, from which wavelets go out exactly as from the original center; and at any instant the new wave front is the envelope of all these wavelets.

Thus, in Fig. 175, if C is the original center of disturbance, and HJK the wave front at a given instant, then the above law affirms that all points (a, b, c, etc.) on HJK may be considered

new centers from which secondary waves go out as from C; and at any later instant the new wave front MNQ is the envelope of the secondary waves, or wavelets, sent out from these centers.

This principle, published by Huyghens in 1690, was one of the most fruitful in the early wave theory. From it many important properties of waves may be deduced.

218. Reflection and refraction. When a system of waves strikes the surface of separation of two media, it is in general broken into two systems, one of which is **reflected** (i.e., turned back) in the first medium. The other passes on into the second medium but is **refracted** (i.e., changed in direction) on passing through the surface. (See Fig. 176.)

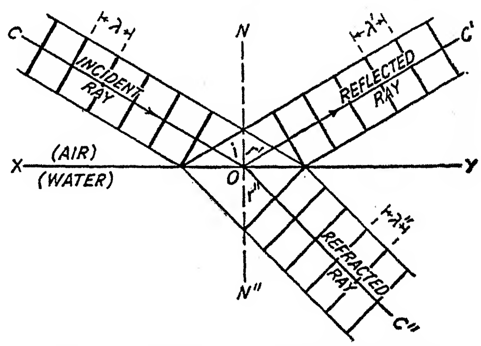


Fig. 176. Reflected and Refracted Beams

In an isotropic medium, i.e., a medium whose physical properties are the same in all directions, a wave of compression or expansion will go outward from a center with the same velocity in all directions; hence the wave fronts will be spherical, and the direction of propagation will be radial, or **normal to the wave front.**

The direction of propagation of a wave is often called a **ray.** From the foregoing paragraph, it follows that in isotropic media the wave front is normal to the ray at a given point.

The **incident ray** is the direction of the wave before striking the surface. The **reflected ray** is the direction of the reflected part of the wave. The **refracted ray** is the direction of the refracted part of the wave.

A **beam** of waves is a portion, or slice, of a wave system comprised within a system of rays, these boundary rays being chosen so as to make the beam convenient for the discussion in which it is used.

The **angle of incidence** i is the angle between the incident ray and the normal to the surface of separation of the two media.

The **angle of reflection** r' is the angle between the reflected ray and the normal to the reflecting surface.

The **angle of refraction** r'' * is the angle between the refracted ray and the normal to the surface of separation of the two media.

In Fig. 176 is shown a beam of waves, say, light waves, striking the surface of separation of air and water. The plane wave fronts are shown perpendicular to the incident ray CO by heavy lines. Similarly, the reflected and refracted wave fronts are shown by lines perpendicular to the reflected and refracted rays, respectively. But these lines are lighter than the incident wave fronts because **the energy of the incident waves is divided between the reflected and refracted wave systems** in a manner depending upon the nature of the media and the surface of separation and upon the angle of incidence.

219. Law of reflection. When a tennis ball rebounds from the ground, or a billiard ball from the cushion, the path of the ball on the rebound makes the same angle with the surface as on approach. It is more satisfactory in scientific work to use the complements of the above angles; i.e., the angles which the paths of approach and recession make with the perpendicular to the surface at the point of impact. These angles also are equal.

The above familiar fact is then equivalent to the *law of reflection:*

1. The angle of incidence is equal to the angle of reflection.

2. The incident ray, the reflected ray, and the normal lie in the same plane.

The first part of this law has been known since the time of Plato (about 350 B.C.). It is readily deduced by geometry with the aid of Huyghens' principle.

In Fig. 177, let CO and OC' be the incident and reflected rays, respectively, for a beam of waves whose advancing wave front is shown by the heavy lines parallel to AE and $A'E'$; and let NO be the normal to XY at O.

If there were only one medium, and hence no surface of separation XY, the waves would proceed parallel to AE; and at a certain time after passing AE, the wave front would have the position $A'E''$.

* The symbol r is used to represent either the angle of reflection or the angle of refraction only when no confusion will occur.

The presence of the surface XY prevents this; but by Huyghens' principle the wavelet from point E on wave front AE proceeds outward in all directions and hence reaches E' in the first medium at the same time that it would have reached E'' if not obstructed. Similarly, the wavelet from D_1 will reach a point D' in the first medium at the same time that it would have reached

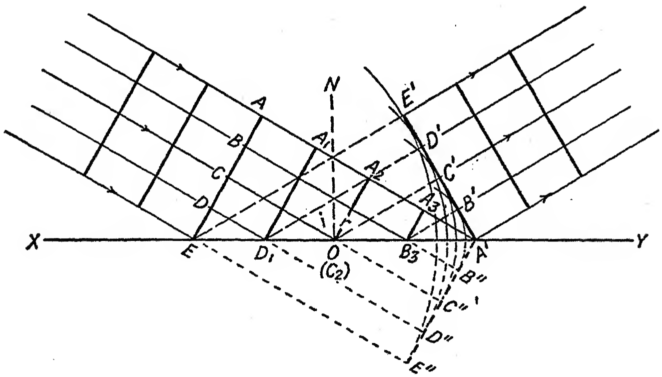


FIG. 177. Reflection by Huyghen's Principle

D'' if unobstructed, and so on for the wavelets from successive points of incidence of wave AE.

Hence, by Huyghens' principle the new wave front is the envelope $A'E'$ of all these secondary waves, or wavelets; and the direction of the reflected beam is OC', perpendicular to $A'E'$.

We then have from geometry,

$$\angle A'E'E \text{ is a right angle}$$

(since a tangent is perpendicular to a radius drawn to the point of contact).

$$EA' = EA' \qquad \text{(being identical)}$$
$$AA' = EE'' = EE' \qquad \text{(by construction).}$$

Therefore,

$$\text{rt.}\Delta\, EAA' = \text{rt.}\Delta\, A'E'E$$

(having the hypotenuse and a side of one equal to the hypotenuse and a side of the other).

$$\angle AA'E = \angle E'EA' \qquad \text{(homologous } \angle\text{s of equal } \Delta\text{s)}$$

or

$$\angle COE = \angle C'OA' \qquad \text{(from similar } \Delta\text{s)}$$

and

$$\angle i = \angle r \tag{188}$$

(complements of equal angles are equal),

which was the relation to be proved.

A beam of waves is said to suffer **regular reflection** if it remains a definite beam after reflection; e.g., light reflected from a polished mirror. If, however, it is **scattered** more or less in all directions by the rugosities of the reflecting surface, it is said to be **irregularly,** or **diffusely, reflected;** e.g., light reflected from white blotting paper.

It may be shown that if regular reflection is to be produced, the height of the irregularities on the reflecting surface may not exceed about 1/16 the wave length λ of the incident wave.*

220. Law of refraction. A stick placed in water and viewed from above appears to be broken where it passes through the surface. This is an example of refraction, a phenomenon observed with waves of all kinds. The law of refraction is best stated in three parts:

1. When a beam of waves passes from one medium into another in which its velocity is less than in the first, it is bent toward the normal, and vice versa.

2. The ratio of the sine of the angle of incidence to the sine of the angle of refraction is a constant for the same two media.

3. The incident ray, the refracted ray, and the normal lie in the same plane.

The first two parts of this law are commonly called Snell's law, after Willebrord Snell, professor of mechanics at Leyden, who discovered it in 1620. It follows easily from Huyghens' principle.

In Fig. 178, let the heavy lines parallel to AE represent portions of an advancing wave front before it strikes the surface of separation XY between two media in which the velocities of the wave are v_1 and v_2, respectively; and let NEN'' be the normal to the surface.

When the wave front strikes XY at E, a wavelet will be sent out from E into the second medium, in accord with Huyghens' principle, which wavelet in 1 sec travels the distance $EE''(= v_2)$, while the wavelet from A travels the distance $AA''(= v_1)$ in the

* R. W. Wood, *Physical Optics* (New York, The Macmillan Company, 1934), p. 43.

first medium. It is assumed that v_1 is greater than v_2. Similarly, the secondary wave from D_1 will reach D'' in the second medium when the one from A reaches A'' in the first; and so on.

The wave front of the refracted wave is then the envelope $A''E''$ of these wavelets. Being a tangent, it is perpendicular to the ray EE'' drawn to the point of contact E''.

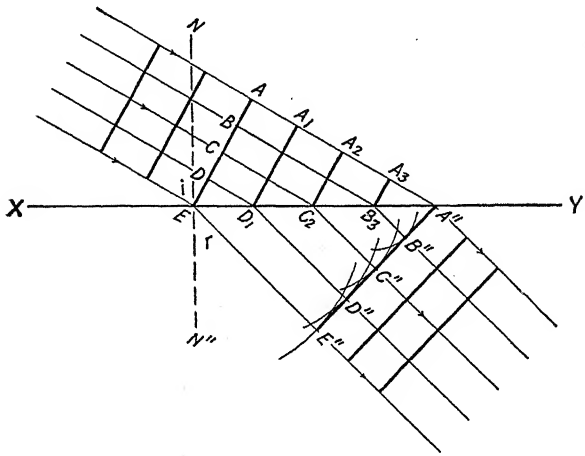


FIG. 178. Refraction by Huyghen's Principle

From the definitions of the trigonometric functions, we then have:

$$\sin AEA'' = \frac{AA''}{EA''} = \frac{v_1}{EA''}$$

$$\sin EA''E'' = \frac{EE''}{EA''} = \frac{v_2}{EA''}.$$

But $\angle AEA'' = \angle i$ and $\angle EA''E'' = \angle r.$ (complements of the same angle are equal)

Therefore $$\sin i = \frac{v_1}{EA''} \tag{a}$$

and $$\sin r = \frac{v_2}{EA''}. \tag{b}$$

Hence, dividing (a) by (b), we have:

$$\frac{\sin i}{\sin r} = \frac{v_1}{v_2} = \text{a constant,} \tag{189}$$

which was to be proved.

In order to secure regular refraction, the height of the irregularities on the surface of separation of the two media must not exceed about 1/4 the length of the incident wave.

221. Index of refraction. The constant of Eq. (189) is called the relative index of refraction of the second medium with respect to the first for the particular kind of wave under consideration. That is,

The index of refraction ($_1\mu_2$) of a second medium relative to a first medium is defined as the ratio of the velocity (v_1) of the wave in the first medium to its velocity (v_2) in the second medium.

Algebraically,

$$_1\mu_2 \equiv \frac{v_1}{v_2}$$

and by Eq. (189),

$$_1\mu_2 = \frac{\sin i}{\sin r}. \tag{190}$$

The value of the relative index of refraction depends upon the kind of wave and the wave length, as well as upon the kinds of media. Thus the mean index of refraction of water relative to air is:

For sound = 0.24
For light = 1.33.

222. Diffraction. It is a matter of common observation that waves which have passed a breakwater AB do not terminate abruptly along the line BC, but bend around the edge of the barrier somewhat as shown in Fig. 179—just as Huyghens' principle would lead us to expect them to do. The phenomenon is common to all kinds of waves and is known as *diffraction*. It was discovered by the Italian physicist, Francesco Mario Grimaldi (1618–1663).

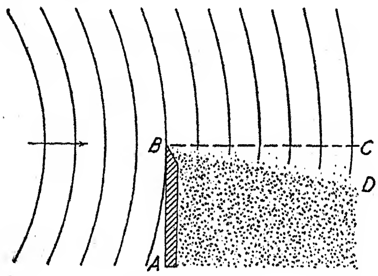


FIG. 179. Diffraction

Diffraction is the bending of waves around the edge of an obstacle.

If diffraction did not occur, the waves would not extend at all beyond the line BC, drawn from the edge of the obstacle in the direction of propagation. BC is called the "edge of the geometric shadow," which is the region screened by the obstacle and extending out to the line BC.

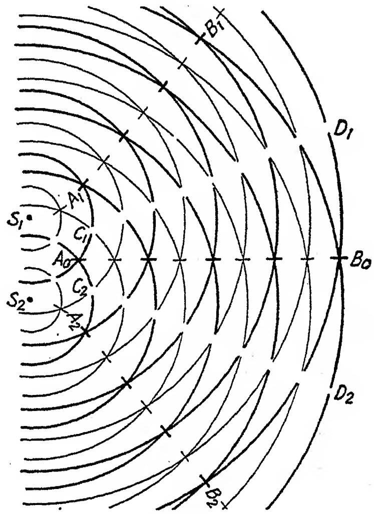


FIG. 180. Interference of Waves

223. Interference. When two trains of waves traverse the same medium at the same time, each particle of the medium executes a motion which is the resultant of the motions it would have in each wave alone.

In Fig. 180, let S_1 and S_2 be the centers of two wave systems, such as would be obtained by moving a two-pronged fork up and down with periodic motion in the surface of a quiet pond. Crests of the two systems are indicated by heavy lines, and troughs, by light lines.

When two waves meet crest on crest or trough on trough, the water is heaped up or depressed to a height equal to the sum of the amplitudes of the two waves. Hence, along the dashed lines A_0B_0, A_1B_1, A_2B_2, we find waves of maximum amplitude. This is called **constructive interference,** or **reinforcement.**

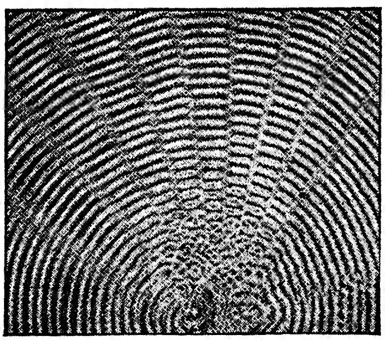


FIG. 181. Photograph of Interfering Waves in Mercury

Where the waves meet crest on trough, however, the resultant amplitude is the difference of the amplitudes of the two waves; and if these amplitudes are equal, the waves destroy each other and no motion of the particles results. This occurs in Fig. 180 along the lines of gaps C_1D_1 and C_2D_2, where the interference is said to be **destructive.**

An actual photograph of interference of two such systems of waves is shown in Fig. 181 (Courtesy of American Book Co.).

224. Stationary waves. A most important case of interference is that which occurs when two like trains of waves traverse the same medium at the same time in opposite directions. In such circumstances we get what are called **stationary,** or **standing,**

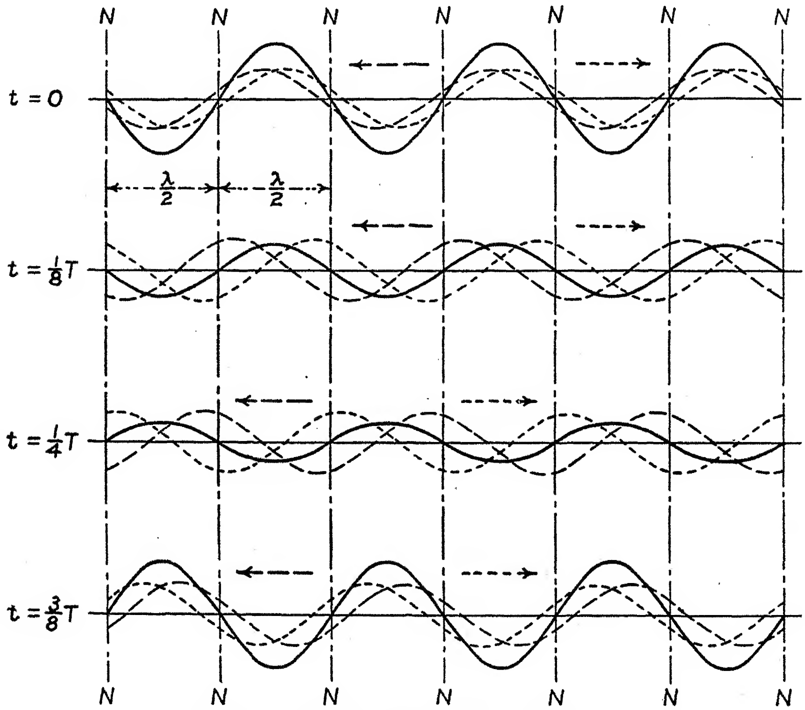


FIG. 182. Construction of Stationary Waves

waves. The necessary conditions are easily secured by reflecting a system of waves back through the same medium.

In Fig. 182, the system indicated by long dashes is moving to the left, while the like system (i.e., having same amplitude and wave length), indicated by short dashes, is moving to the right.

At a given instant ($t = 0$), they have the positions shown. Each particle of the medium, say, a string, will have at any moment a displacement equal to the algebraic sum of the displacements due to the two waves separately. The resulting positions of the particles, or of the string, when $t = 0$, are shown by the heavy line.

Other positions of the component waves and of their resultant wave are shown for the instants indicated at the left. It will be observed that the particles whose two displacements neutralize each other occur always at the same points of the string, which therefore lie on the vertical lines *NN*.

These points where the motion is zero (or a minimum) are called **nodes** (= knots).

All particles of the medium between **nodes** oscillate with simple harmonic motion, but with different amplitudes. The amplitudes

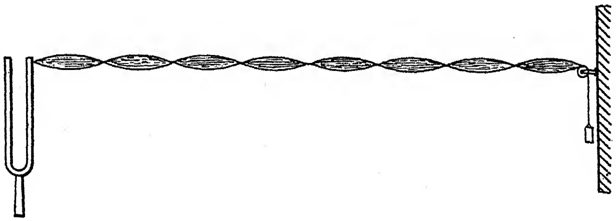


FIG. 183. Melde's Experiment

of the particles increase from zero at the nodes to a maximum at points midway between the nodes.

These points of maximum amplitude are called **antinodes.**

When stationary waves occur in a cord at a frequency greater than about 16 v.p.s., the eye cannot distinguish the cord in its successive positions, but sees a hazy spindle-like figure, or loop, between each pair of nodes, as in Fig. 183.

It is because the wave forms appear to stand still that waves of this kind are known as **stationary,** or **standing, waves;** but it must be clearly understood that the particles of the medium vibrate, except those at the nodes.

From the figure, the distance between nodes is seen to be one-half the wave length ($\frac{1}{2}\lambda$); and it may easily be shown that in order to secure stationary waves in a string, the following relation must be fulfilled:

$$\frac{2l}{v} = kT \tag{191}$$

where l is the length of the string;
v is the velocity of the wave;
T is the period of the vibration of particles; and
k is any whole number—1, 2, 3, 4, etc.

Stationary waves are of frequent occurrence in nature. In sound they occur whenever there is a sustained tone, as from a piano string or an organ pipe.

225. Polarization of waves. A wave whose particles all vibrate in parallel planes is said to be **plane-polarized.** Such waves are obtained in a string one end of which is fastened and the other moved up and down in a straight line with periodic motion.

If, however, the free end of the string is moved uniformly around a circle whose plane is perpendicular to the length of the

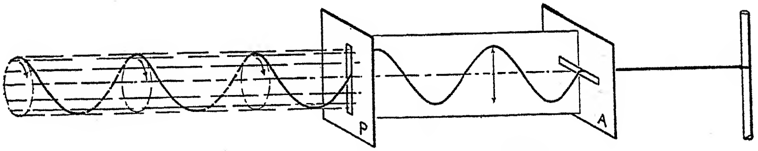


FIG. 184. Circularly Polarized Waves

string, the string will take the form of a helix, or screw thread, and the waves are said to be **circularly polarized.**

Transverse waves can always be polarized, for some type of screen or other filter can be found to stop vibrations in one plane and pass those in another plane, usually at right angles to the first. Thus, in Fig. 184, a circularly polarized wave in a string is plane-polarized by being passed through a straight slot in a board *P*. Vibrations perpendicular to the slot are stopped; only

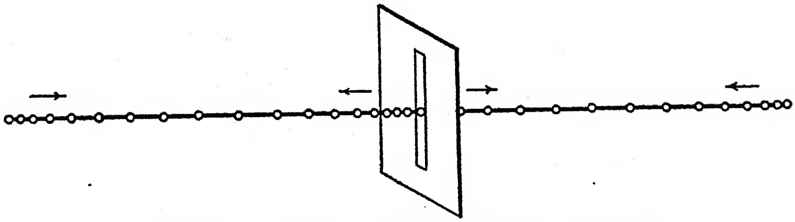


FIG. 185. Longitudinal Waves Cannot Be Polarized

those parallel to its sides are passed. Hence the emerging wave is **plane-polarized.** If the string passes on through another similar board *A* with its slot at right angles to the first one, the **plane-polarized** wave is completely stopped, so that there is no wave motion in the string beyond the second board.

The first device *P* which produces polarization is called the

polarizer; the second A, which is used to determine whether or not the wave is polarized, is called the **analyzer.**

No such process is possible with longitudinal waves, however, because their particles vibrate along the direction of propagation, and hence the wave comes through unaltered (Fig. 185).

Polarization is therefore considered the test of the kind of wave.

1. **If a wave can be polarized, it is transverse.**
2. **If a wave cannot be polarized, it is longitudinal.**

226. Intensity of waves. The intensity E of a wave is defined as the amount of energy conveyed by the wave per second through unit area taken at right angles to the direction of propagation.

Consider a thin prism of the medium having an area of 1 cm² at right angles to the direction of propagation, and having a thickness t parallel to that direction and so small that all the particles of the prism have sensibly the same amplitude of vibration r. Let the density of the medium be D. Then,

$$\text{Volume of prism} = 1\ \text{cm}^2 \times t\ \text{cm} = t\ \text{cm}^3$$

and

$$\text{Mass of prism} = Dt\ \text{gm}.$$

Assuming the particles of the medium to have shm, as they usually do, the velocity of each particle at the middle of its path is the same as the velocity V of the auxiliary point Q (see Sec. 35). That is,

$$V = \frac{2\pi r}{T} = 2\pi rn$$

where T and n are the period and frequency, respectively, of the vibration.

Since each of its particles has this maximum velocity V at a certain instant, the kinetic energy of the medium within the prism (which at that instant is its total energy) is:

$$\begin{aligned}\text{Energy} &= \tfrac{1}{2}(Dt)V^2 = \tfrac{1}{2}(Dt)(2\pi rn)^2 \\ &= 2\pi^2 Dtr^2n^2 \quad \text{ergs}\end{aligned}$$

and the energy per unit of volume is the

$$\text{Energy density} = \frac{2\pi^2 Dtr^2n^2\ \text{ergs}}{t\ \text{cm}^3} = 2\pi^2 Dr^2n^2 \quad \text{ergs/cm}^3.$$

If the velocity of propagation of the wave is $\mathfrak{v}$ cm/sec, the energy

transmitted per sec through the 1 cm² in question is the same as that of v cm³ of the medium. Hence we have:

$$\text{Intensity of wave} = \left(v\,\frac{\text{cm}^3}{\text{cm}^2\text{-sec}}\right)\left(2\pi^2 D r^2 n^2\,\frac{\text{ergs}}{\text{cm}^3}\right)$$

$$= 2\pi^2 v D r^2 n^2\,\frac{\text{ergs}}{\text{cm}^2\text{-sec}}. \qquad (192)$$

Thus it is seen that, other factors being constant, **wave intensity varies directly as the square of the amplitude r and the square of the frequency n.**

227. Damped and undamped waves. If a particle is vibrating in a resisting medium, like a pendulum in molasses, it does work

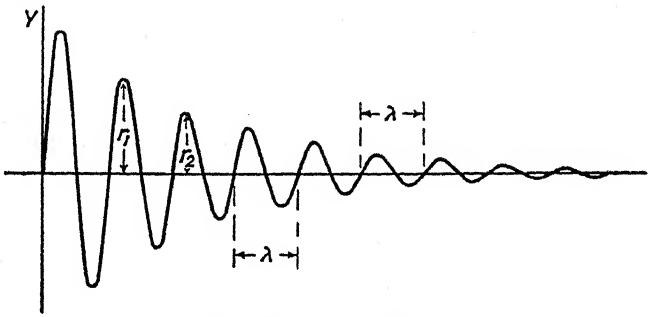


FIG. 186. Damped Wave

on the medium; consequently its energy and amplitude gradually decrease. Most media offer some resistance to the motion of their particles, or to the change of condition, that constitutes a wave. In such media, the amplitude of a wave at a given point will gradually diminish unless additional energy is supplied; and waves receiving energy from a source of constant power will diminish in amplitude as they recede, on account of the work done in overcoming the resistance of the medium.

Such waves are called *damped waves* and are illustrated by Fig. 186. In damped waves, the ratio of any amplitude r_1 to the next succeeding amplitude r_2 of the same sign is a constant; and it may be shown that

$$\frac{r_1}{r_2} = \text{a constant} = \epsilon^{\alpha T} \qquad (193)$$

where T is the period and α is the "damping factor."

Taking the logarithm of Eq. (193) to the base ϵ, we have

$$\delta \equiv \log_\epsilon \frac{r_1}{r_2} = \alpha T$$

where δ is called the **logarithmic decrement.**

If the loss of energy due to the resistance of the medium is negligible, the wave is **undamped** and the energy of a spherical wave within a spherical shell one wave length thick will remain constant (Fig. 187); that is the total energy W entering two such shells of inner radii R_1 and R_2, respectively, in 1 sec is:

$$W = 4\pi R_1^2 E_1 = 4\pi R_2^2 E_2$$

whence

$$\frac{E_1}{E_2} = \frac{R_2^2}{R_1^2}. \qquad (194)$$

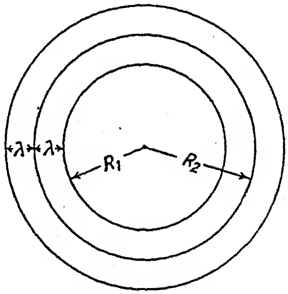


FIG. 187

Hence, **in spherical waves, when damping is negligible, the intensity is inversely proportional to the square of the distance from the source.** And on combining Eqs. (192) and (194), we find that **the amplitude varies inversely as the distance from the source.**

PROBLEMS

1. A siren disk (see Fig. 197) has 64 holes and rotates 1500 times per minute. What is the pitch (frequency) of the sound vibrations?

2. What is the wave length of the above sound waves when the temperature of the air is 20°C?

3. If a circular saw has 60 teeth and produces a note of 2000 vibrations per second what is the speed of the saw?

4. What is the wave length of the sound waves in the preceding problem if the velocity of sound in air is then 1120 ft/sec?

5. If Young's modulus for brass is $9.2 \times 10''$ dynes/cm², find the velocity of sound waves in brass.

6. If the velocity of sound in steel is 5000 m/sec, what is Young's modulus for steel?

7. An increase of pressure of 1 atm reduces the volume of water by 0.000050 of the original volume. Find the velocity of sound in water.

8. If the adiabatic coefficient of elasticity of air is 1.42×10^6 dynes/cm², compute the velocity of sound in air at 0°C.

CHAPTER XVII

SOUND

228. The nature of sound. Galileo (1564–1642), while scraping a metal plate with a chisel, observed that when a clear note was produced, the plate was marked by indentations equally spaced. This was construed to indicate that sound was due to vibrations of the plate produced in the scraping process.

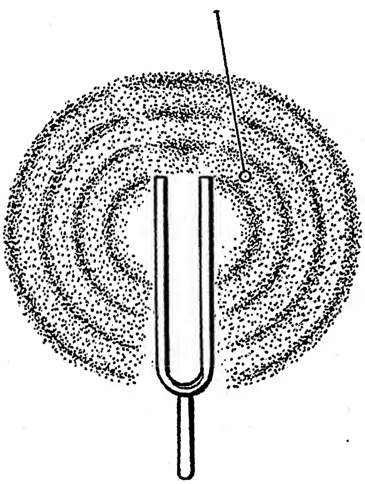


FIG. 188. Sound Waves from Tuning Fork

All subsequent experiments confirm this fact that sound originates from vibrating bodies. A pith ball suspended by a thread and placed against a tuning fork, as in Fig. 188 shows by its violent agitation that the fork is in vibration; and manometric flames attached to an organ pipe are caused to flicker regularly when the pipe is sounding. But the most convincing evidence is that of actual photographs of sound waves taken by Professor A. L. Foley of Indiana University, which show clearly the **compression and rarefaction which constitute the wave** (Fig. 189).

We therefore make the following definition: **Sound consists of compressional waves capable of affecting the auditory nerve so as to produce the sensation of hearing.**

Only ponderable (i.e., having mass) elastic media transmit sound. A bell in a vacuum is not heard, for though a vacuum will transmit light, it contains no medium having mass to transmit sound. Snow and a feather pillow transmit sound very poorly: they have mass, but very little elasticity.

For the most part, we deal with sound in air; but water, metals, wood, and most common substances transmit it well.

Sound-deadening substances are inelastic and of loose, porous texture such as felt, rock wool, seaweed, and board fabricated from wood pulp and cane stalks. The energy of the sound waves is dissipated by repeated reflections within the heterogeneous texture of these materials.

229. Speed of sound. As determined by experiment, the speed of sound is in good agreement with the values predicted by Newton's formula, Eq. (184). At first this appeared not to hold true for gases, but Laplace suggested that the gases were heated during compression and cooled during expansion so rapidly that isothermal conditions were impossible; and that, therefore, the **adiabatic** and not the isothermal coefficient of elasticity should be used.

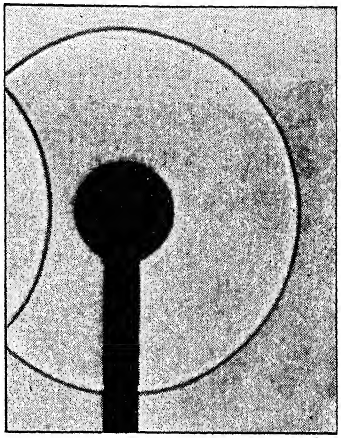


FIG. 189. Reflection of Sound Wave from Plane Surfaces. (Courtesy of Professor A. L. Foley)

The adiabatic coefficient may be shown to be γ times the isothermal coefficient, where γ is the ratio of the specific heat of the gas at constant pressure to the specific heat at constant volume ($\gamma = 1.40$ for air). That is, for gases,

$$v = \sqrt{\frac{\gamma E}{D}}.$$

But since the isothermal coefficient of elasticity of a gas is equal to its pressure P,* this may be written:

$$v = \sqrt{\frac{\gamma P}{D}} \tag{195}$$

and the values from this equation check with experiment.

Probably the best value for the speed of sound in free air is that obtained by D. C. Miller, Professor of Physics at the Case School of Applied Science, from observations made at Sandy Hook in

* W. Watson, *op. cit.*, p. 365.

1934, using coast defense guns as sources. His value is 1087.13 ft/sec (= 331.36 m/sec) at 0°C.

The following table gives the speeds of sound in some other common substances.

SPEEDS OF SOUND
From Kaye and Laby's Tables

Medium	Speed
Hydrogen (std. conditions)	1286 m/sec
Oxygen (" ")	317.2
Sulphur dioxide	209
Water (at 13°C)	1437
Brass	3650
Steel	4950
Lead	1250
Oak	4200

230. Variation of speed of sound:

1. *With temperature.* From the general gas law (see Sec. 265),

$$PV = RMT$$

where R is the gas constant and T is the absolute temperature. Then

$$P = R\frac{M}{V}T = RDT.$$

Therefore

$$\frac{P}{D} = RT. \tag{a}$$

Substituting this value in Eq. (195),

$$v = \sqrt{\gamma RT}$$

where $\sqrt{\gamma R}$ is constant ($\equiv c$),

so that

$$v = c\sqrt{T}. \tag{b}$$

Hence, **the speed of sound in a gas is proportional to the square root of its absolute temperature**, which may be written:

$$\frac{v}{v_0} = \frac{\sqrt{T}}{\sqrt{T_0}}$$

where v_0 is the speed at 0°C, and T_0 is the corresponding absolute temperature.

Substituting for T and T_0 their values, we get the more usual form:

$$v = v_0\sqrt{\frac{273 + t}{273}} = v_0\sqrt{1 + \frac{1}{273}t}. \tag{196}$$

If the temperature rises 1°C, we find, on making $v_0 = 331.36$ m/sec and $t = 1°$ in this equation, that $v = 331.96$ m/sec. That is, the speed of sound in air increases about 0.6 m/sec, or about 2 ft/sec, per 1°C rise of temperature. This value may be used for approximate computations.

2. *With pressure.* From Eq. (a) above,

$$\frac{P}{D} = RT.$$

Since R is constant, P/D is constant when T is constant; that is, when P is doubled, D is doubled, etc. Hence it follows that **the speed of sound in a gas, which is $\sqrt{\gamma(P/D)}$, is unaffected by changes in average pressure if the temperature is constant.**

3. *With molecular velocities.* Since a gas cannot sustain a tensile stress, sound must be transmitted through a gas by molecular collisions. We should therefore expect a definite relation between the speed of sound and the effective speed * of the molecules of the gas.

In Sec. 269 we shall see that the absolute temperature T of a gas is proportional to the average kinetic energy of its molecules provided the temperature is not very near the absolute zero:

$$\begin{aligned} T &\propto \overline{E} \\ T &= c'\overline{E} = c'(\tfrac{1}{2}m\overline{v}^2) \\ &= c''\overline{v}^2 \end{aligned}$$

where

$$c'' \equiv \tfrac{1}{2}c'm.$$

Substituting this for T in Eq. (b) above,

$$v = c\sqrt{c''\overline{v}^2} = c'''\sqrt{\overline{v}^2}$$

* Effective speed is the square root of the average squared speed.

$$c''' \equiv c\sqrt{c''}$$

where

$$v = c'''\bar{v}$$

or

$$v \propto \bar{v}. \tag{c}$$

Hence, the speed of sound in a gas is proportional to the effective speed of the molecules of the gas.

4. *With intensity.* By experiments made near Versailles in 1864, Henri Regnault (1810–1887) showed that sounds of great intensity travel with greater speeds than do those of low intensity. The difference is small, however, being of the order of 1/3 of 1% in these experiments.*

5. *With frequency.* In 1905, Violle and Vautier found that the speed of sound is independent of frequency (pitch). On this account the music of a band is as pleasing at a considerable distance as near by, for the notes of all the instruments retain their proper time, pitch, and quality unchanged. However, notes of high frequency are more attenuated (weakened) by distance than are those of low frequency.†

231. Determination of the speed of sound.

(a) The most obvious method of determining the speed of sound is to observe the interval between seeing the flash of a gun and hearing its report at a known distance. Light has the enormous speed of 186,000 mi/sec, and hence the time for it to travel the distance may be neglected. This method, with refinements, was used by Regnault in the huge water mains of Paris, where the air was quiet and of uniform temperature.

(b) The method of resonance is usually employed in the laboratory.

Resonance is the reinforcement of one series of vibrations by another series which has the same frequency (cf. Sec. 152).

For the determination of the speed of sound in air, a tuning fork F of known frequency n is mounted just above a long vertical tube H in which the depth of the water may be varied by raising or lowering the tank J (Fig. 190).

The fork is set in vibration and the water adjusted until the

* Barton, *Text-book on Sound* (New York, The Macmillan Company, 1908), p. 516.
† *Ibid.*, p. 523.

sound is of maximum intensity. The length L of the tube from its open end to the water surface is then approximately $1/4\lambda$. This may be shown as follows:

When the prong of the fork is in position 1, it starts a compression down the tube. This compression is reflected at the surface of the water, and in order to reinforce the next upward compression, it must return to the mouth of the tube just as the fork starts from position 2 on its upward trip.

Hence, while the fork makes half a vibration, from 1 to 2, a compression must travel down the tube and back again. That is, in order to secure resonance, the sound wave must travel $2L$ while the fork makes one-half a vibration; and therefore it must travel $4L$ while the fork makes a complete vibration. But the distance a wave travels while the source makes a complete vibration has been shown to be the wave length (Sec. 215).

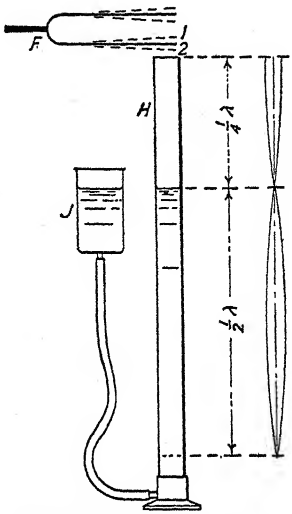


FIG. 190. Resonance Tube

Hence,

$$\lambda = 4L \text{ (approximately)}.$$

Lord Rayleigh has shown * that a correction is necessary on account of the end effect of the tube, so that, more accurately,

$$\lambda = 4(L + .6r) \tag{197}$$

where r is the radius of the tube.

Since the air particles in contact with the water are necessarily at rest, there is a **node** at the water surface. Other lengths that give resonance will be found as the water is lowered. Each of these positions of the surface is also a node; and we have in the tube a system of stationary waves, which occur whenever there is a sustained tone. The distance between nodes is $1/2\lambda$ by Sec. 224. Hence we get resonance for lengths of the tube as follows:

$$\tfrac{1}{4}\lambda,\ \tfrac{3}{4}\lambda,\ \tfrac{5}{4}\lambda,\ \tfrac{7}{4}\lambda, \text{ etc.}$$

* *Ibid.*, p. 553.

After λ has been found, we obtain the velocity at once by Eq. (182):

$$v = n\lambda.$$

(c) The *Kundt's tube method*, devised by Professor Kundt of Berlin in 1865, is the one generally used for determining the relative speeds of sound in different gases and its speed in solids.

As is shown in Fig. 191, a glass tube G, about 1 meter long and 4 or 5 cm in diameter, is tightly closed at one end by a stopper.

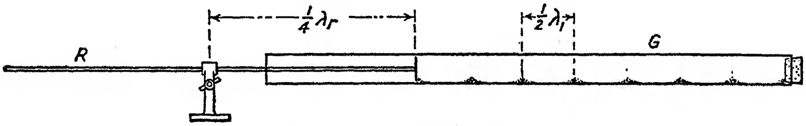


FIG. 191. Kundt's Tube

In the other end is inserted a rod R of metal or glass, clamped at its center and carrying a light cardboard piston that fits in G with just enough clearance to move freely.

Cork dust, lycopodium, or other light powder is distributed in a small uniform ridge along the bottom of G. The rod is made to vibrate longitudinally by grasping it firmly with a cloth on which rosin has been sprinkled freely, and giving it a single vigorous stroke from the clamp to the left end.

An intense musical note will be emitted by the rod, and if the length of the tube is properly adjusted, stationary waves will be set up in it. The dust is blown aside at the antinodes and remains in little heaps at the nodes, where the motion of the particles of the gas is zero. The stopped end is necessarily a node, and the piston end is nearly a node, since the actual motion of the end of the rod is small. The length of the tube is therefore divided into a whole number of divisions, each of length $1/2\lambda$.

By using the same rod, we keep n constant so that the velocities v_1 and v_2 in different gases are, respectively,

$$v_1 = n\lambda_1$$

$$v_2 = n\lambda_2$$

and

$$\frac{v_1}{v_2} = \frac{\lambda_1}{\lambda_2}. \qquad (198)$$

That is, the velocities are proportional to the distances between the nodes for different gases.

Also, the center of the rod is a node (because it is clamped), and the free ends are antinodes. Hence, the distance from the clamp to either end is $1/4\lambda_r$, where λ_r is the wave length in the rod.

When the speed in air is known, Eq. (182) may be solved for the frequency n; and having that, we are able to compute the speed of sound in the material of the rod by the relation,

$$v_r = n\lambda_r. \tag{199}$$

232. Reflection of sound. Like all other types of waves, sound is reflected—at least in part—whenever it strikes the surface of separation between two media. The most familiar example of the reflection of sound is the **echo.** A wall or a cliff, whose irregularities are small compared to the wave length of the sound, will return a sound which may be clearly distinguished from the original, provided the source is at a sufficient distance. The distance necessary may be computed as follows.

It is known from experiment that two sounds of the same kind cannot be distinguished if the interval between their occurrences is less than 1/10 sec. In 1/10 sec sound will travel 108.7 ft in air at 0°C. Hence the distance from the source to the reflecting surface and back again must be at least 108.7 ft, and **the distance between source and reflecting surface must be at least**

$$\tfrac{1}{2} \text{ of } 108.75 = 54.4 \text{ ft (approximately)}$$

in order that the reflected sound may be distinguishable from the original and consequently recognizable as an echo.

If a watch S (Fig. 192) is placed in front of a smooth, concave, spherical surface of wood or metal, but not too close, its ticking will be brought to a focus at some point S', in a manner similar to the behavior of light (see Sec. 526). In fact, the same mirror may be used for light and for sound, and the same points will be found to be conjugate foci for both.

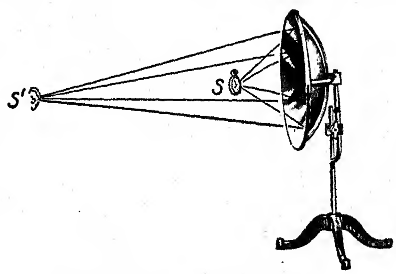


FIG. 192. Sound Reflector

Before the advent of modern electrical public address apparatus, it was customary to place a sounding board (i.e., a

sound reflector) behind the speaker's stand in churches and auditoriums. There are many "whispering galleries," notably in the Capitol at Washington and in the Mormon Tabernacle at Salt Lake City. When one stands at a certain spot in one of these buildings, he can distinctly hear a whisper uttered at a certain other spot on the opposite side of the room, though it is inaudible at intermediate points. The two points are conjugate foci for sound (see Sec. 545).

Figure 189 shows a sound wave being reflected from a plane surface.

Reflection of sound is employed in a device called a **fathometer** for determining depths at sea. An intense beam of sound waves is sent out by an oscillator located at the bottom of the ship. The reflected wave, on its return, actuates a device similar to a telephone receiver, whose signal is amplified and flashes a neon light carried on a revolving hand. When the hand passes zero the signal is sent, and on its return the neon light flashes. The depth is read directly on the dial at the position of the hand when the flash occurs (cf. Sec. 253).

A somewhat similar method is used for locating oil. A heavy charge of explosive is detonated at various points in a region where oil is thought to exist. The time and intensities of the returning waves are automatically recorded at a chain of stations, and from these data the location of salt domes (characteristic of oil regions) can be predicted with some assurance.

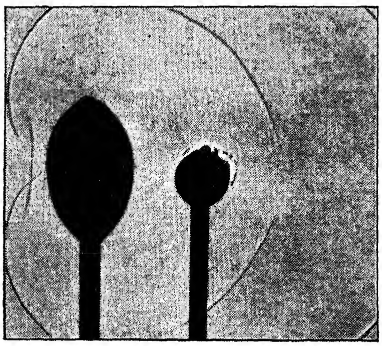


Fig. 193. Refraction of Sound by Lens. (Courtesy of Professor A. L. Foley)

233. Refraction of sound, though less noticeable than reflection, occurs whenever a beam of sound passes from one medium into another in which its velocity is different from that in the first medium.

Figure 193 shows a photograph of a sound wave that has been changed by refraction through a sound lens, consisting of a lens-shaped bag filled with SO_2 in which its speed is less than in air. The original spherical wave is seen

at the extreme right; the refracted plane wave at the left of the lens. The source of the sound is an electric spark produced behind the circular electrode in the center of the figure.

If the lens had been filled with hydrogen, in which the speed of sound is greater than in air, the wave after passing through the lens would still have been spherical, but its radius would have been less than before striking the lens.

On account of passing through bodies of air of different densities, sounds such as the firing of guns are sometimes heard at a considerable distance when not audible at nearer points. The same effect may be due to the change in the direction of the sound waves by the wind. For these reasons, also, fog signals may pass completely over a ship well within their ordinary range.

234. Interference of sound: Beats. If two sources of sound of the same frequency, such as two tuning forks A and B, are sounded at the same time, the resulting sound is smooth and the two sources are said to be in *unison*, or resonance.

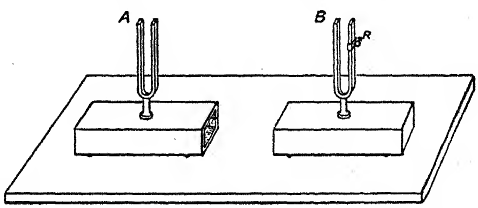


Fig. 194. Tuning Forks Producing Beats

But if the frequency of one of the forks is changed (lowered), as by the addition of a rider R (Fig. 194), a throbbing effect called *beats* is produced when the forks are sounded together.

Beats are regular outbursts of sound followed by periods of comparative quiet, caused by the interference of two systems of

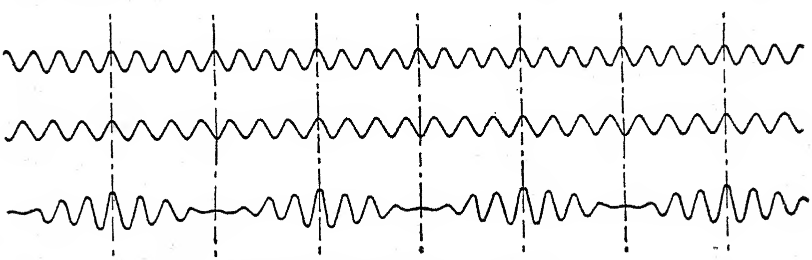


Fig. 195. Wave Form of Beats

waves of different frequencies. The phenomenon of beats is readily understood from Fig. 195, which may represent two wave

trains of any kind (by the convention of Sec. 214), having frequencies of 24 and 21 vps (vibrations per second), respectively.

Where crest falls on crest, we have constructive interference: the amplitude, and consequently the intensity, is a maximum (Sec. 226). Where crest falls on trough, there is destructive interference: the opposite forces nullify each other and the amplitude and intensity are minimum.

It will be seen that a crest of the wave of greater frequency overtakes a crest of the one of lower frequency a number of times in one second equal to the difference of the two frequencies—in this case, three times.

Hence, **the number of beats per second is equal to the difference of the frequencies.**

If the number of beats per second is greater than about 30, they produce a new note called the **beat note.** Beats are used in the tuning of musical instruments. In a piano, for example, the strings of middle C are first tuned until they give no beats with a Standard Middle C tuning fork. Other notes are then tuned until they produce the proper numbers of beats per second with those already tuned. The heterodyne type of radio receiving set employs the beat note produced by the incoming waves and waves generated in the receiver itself; and the very vexing "whistling" often heard is the beat note resulting from the interference of the broadcast wave with the wave produced by a local regenerative receiving set.

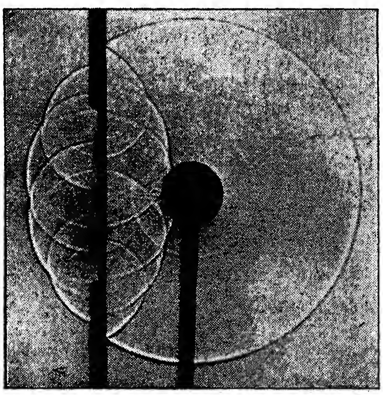


FIG. 196. Sound Waves Through Grating. (Courtesy of Professor A. L. Foley)

235. Huyghens' principle (Sec. 217) for sound is beautifully verified by another of Professor Foley's photographs, shown in Fig. 196. The compression and rarefaction composing the original wave are shown by the dark circle and its accompanying light circle at the right of the figure, their centers being behind the central electrode.

At the left, this spherical wave impinges upon a "grating"

consisting of a screen pierced by five horizontal slots perpendicular to the paper. Beyond the grating on the left are seen the transmitted wavelets, with centers at the centers of the slots; and to the right of it are the reflected wavelets, with centers at the centers of the strips.

It will be noted that the original wave front, if continued, would neatly envelop the transmitted wavelets. The envelope of the reflected wavelets would then be a new wave front whose virtual center is as far behind the reflecting surface as the center of the original wave was in front of it.

236. Musical sounds and noise. We consider sounds to be musical that are smooth and pleasing to the ear; whereas harsh, jangling sounds that are disagreeable to the ear are called noises. Upon examination of these two classes of sounds it is found that: **Musical sounds are due to regular vibrations; noises are due to irregular vibrations.**

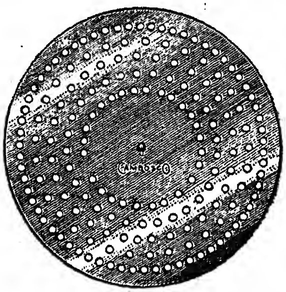


FIG. 197. Siren Disk

This is well illustrated by the siren. The siren disk (Fig. 197) is pierced by five circular rows of holes. The holes of the four outer circles are equally spaced, but those of the innermost circle are irregularly spaced. When this disk is mounted on a rotator and turned rapidly, a blast of air directed at any one of the four outer circles yields a **musical note** because the puffs that pass the holes come through at regular intervals. But when one blows against the innermost circle of holes, the sound is a rough, unpleasant **noise,** for the irregularly spaced holes produce irregularly timed puffs, or sound waves.

237. Characteristics of musical sounds. Waves in general are characterized by frequency, amplitude, and wave form. It is to be expected that these properties of the wave would produce corresponding properties of sound, as they do.

The characteristic properties of musical sounds are pitch, loudness, and quality.

Pitch depends upon the frequency of the sound wave. It may be roughly described as the shrillness of the sound.

Loudness of sound is a measure of sensation, which belongs in the province of psychology. It depends upon intensity and frequency, and intensity has been shown (Sec. 226) to be proportional to the square of the amplitude and the square of the frequency. Hence, **loudness depends upon the amplitude and the frequency of the sound wave,** but the relation is not definitely known.

Lord Rayleigh has found that the amplitudes of sound waves in air vary from about 10^{-6} mm for sounds that are just audible to 1 mm for the loudest sounds that can be heard.

Quality depends upon the wave form. It is the property which enables us to distinguish the notes of a piano from those of a violin or an organ.

The first two of these characteristics of sound may be demonstrated by means of the siren of Fig. 197. As each hole comes opposite the nozzle, a puff of air passes through the hole and produces a sound wave on the far side of the disk. Accordingly, as many waves are sent out per second as holes pass the nozzle per second. There are more holes around the larger circles, hence the frequency of the sound is greater when the blast is directed at the outer circles of holes.

As the blast is moved to the smaller circles of holes, there is a noticeable change in the character of the sound. That from the larger circles is said to be of **higher pitch** (greater frequency) than that from the smaller circles. Hence the pitch depends upon the frequency.

To produce louder sounds one must blow harder, i.e., impart more energy to the sound. The harder one blows, the farther each puff will extend beyond the passing hole, and, therefore, the greater the amplitude. Thus, loudness depends upon the amplitude of the waves.

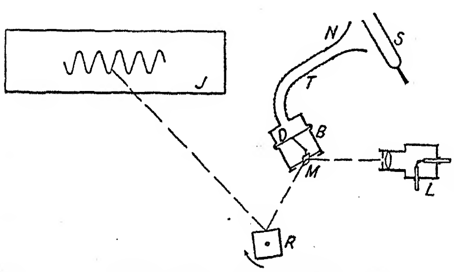


FIG. 198. Phonodeik

The fact that the quality of a musical sound depends upon its wave form is demonstrated by examining the wave forms of notes from various musical instruments by means of some type of oscillograph. One of the best of these is called

the **phonodeik** and was devised by D. C. Miller. Its principle is shown in Fig. 198.

The sound wave produced by the source S enters the mouthpiece N and is transmitted through the flexible tube T to the light diaphragm D, which completely closes the cross section of the circular box B. Each compression deflects D outward, and each rarefaction permits it to spring back. These vibrations are communicated to a tiny mirror M, mounted on a delicate metal ribbon placed horizontally across the diameter of the box, and they cause M to execute small oscillations about this ribbon as an axis.

A narrow beam of light from a strong source L is reflected from M to a four-sided mirror R rotating about a vertical axis. The mirrors R reflect the beam of light to the screen J, where it makes a bright spot. The motion of the mirror M gives this spot vertical displacements, while the mirrors R give it horizontal

FIG. 199. Wave Forms

displacements. The spot, therefore, traces out on J a curve showing the variations of pressure of the wave on D; that is, the curve represents the wave form of the sound.

Fig. 199 shows characteristic wave forms for the tuning fork and for certain notes of the oboe, the violin, and the human voice.

The tuning fork is seen to give almost a pure sine wave. The irregularities, or departure from sine waves, of the other forms are due to the overtones of the respective instruments.

If the screen J (Fig. 198) is replaced by a photographic film, permanent records of the curves are obtained. These curves may be resolved into component curves by means of an apparatus called a *harmonic analyzer*.* In all cases it develops that the necessary component curves are harmonic, or sine, curves of various amplitudes, frequencies, and phases.

238. Fourier's theorem. The above result was predicted in 1822 by the French physicist Joseph Fourier, who showed mathe-

* D. C. Miller, *Science of Musical Sounds* (New York, The Macmillan Company, 1916), p. 97.

matically that any persistent periodic curve, $y = f(\omega t)$, could be expressed as a series of sine or cosine functions.* Thus,

$$y = y_1 + y_2 + y_3 + \cdots + y_n$$

where

$$y_1 = r_1 \sin(\omega t + \alpha_1) \qquad y_3 = r_3 \sin(3\omega t + \alpha_3)$$
$$y_2 = r_2 \sin(2\omega t + \alpha_2) \qquad y_n = r_n \sin(n\omega t + \alpha_n)$$

or

$$y = r_1 \sin(\omega t + \alpha_1) + r_2 \sin(2\omega t + \alpha_2) + \cdots + r_n \sin(n\omega t + \alpha_n). \quad (200)$$

Each of the components of y, i.e., $y_1, y_2, y_3, \ldots y_n$, is a sinusoid, or harmonic curve, and hence they are called the **harmonics** of the original complex curve, $y = f(\omega t)$. The r's are the amplitudes, and the α's are the epoch angles and determine the phases when $t = 0$.

By Eq. (43), the periods of these harmonic components are:

$$\frac{2\pi}{\omega}, \frac{2\pi}{2\omega}, \frac{2\pi}{3\omega}, \cdots \frac{2\pi}{n\omega}$$

and are seen to be proportional to the numbers

$$1, \frac{1}{2}, \frac{1}{3}, \cdots \frac{1}{n}.$$

The wave lengths are likewise proportional to these numbers; and the frequencies, being reciprocals of the periods, are proportional to the numbers,

$$1, 2, 3, \ldots n.$$

The component curve, or function, having the lowest frequency is called the **fundamental, or first, harmonic**; the one having double this frequency, the **second harmonic**; and so on, the order of the harmonic being the same as the coefficient of ω.

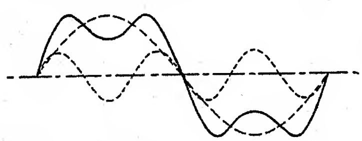


FIG. 200. Composite Wave

Figure 200 shows the wave form obtained by combining as components a fundamental and a third harmonic, the components being shown by dashes.

* *Ibid.*, p. 92.

239. Doppler's principle in sound. The Doppler effect is most often observed in sound waves, although it may occur in any type of wave motion. If one is standing near a railway when a whistling locomotive passes, the pitch of the whistle is noticeably lower as the engine recedes than when it is approaching. An organ reed mounted at the end of a horizontal bar and rotated rapidly about a vertical axis produces distinct beats, since the pitch is higher (wave length shorter) while it is moving toward the observer than when moving away from him.

In any such case, the **actual wave length λ** and the **apparent frequency n'** are readily computed.

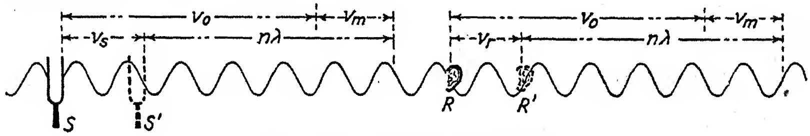


FIG. 201. Doppler's Principle

In Fig. 201, let

S be a source of sound moving to the right with speed v_s;
R be a receiver of sound moving to the right with speed v_r;
v_m be the speed of the medium to the right;
v_0 be the speed of sound in the medium at rest;
λ be the actual wave length;
n be the actual frequency of the source; and
n' be the apparent frequency of the source as observed at the receiver.

Let all speeds be relative to the earth and positive, positive direction being taken from the source toward the receiver.

In 1 sec a given wave front leaving the fork would move forward a distance v_0 in still air. If the air has a speed v_m, it will move this additional distance; so the total distance it will move in 1 sec is $(v_0 + v_m)$. At the same time the source will advance a distance v_s; hence the n waves sent out by the source in this second are comprised in the distance $(v_0 + v_m - v_s)$. That is,

$$n\lambda = v_0 + v_m - v_s$$

$$\lambda = \frac{v_0 + v_m - v_s}{n}. \tag{201}$$

If the receiver R were at rest, a given wave front in 1 sec would

pass him a distance $(v_0 + v_m)$. But since he moves in 1 sec a distance v_r in the same direction, this wave front passes him the distance $(v_0 + v_m - v_r)$ only. The number of waves that pass him (strike his ear) in 1 sec is the *apparent* frequency n' of the source, and these are comprised in the distance $(v_0 + v_m - v_r)$. Hence,

$$n'\lambda = v_0 + v_m - v_r$$

$$n' = \frac{v_0 + v_m - v_r}{\lambda}$$

in which substituting the value of λ gives:

$$n' = \frac{v_0 + v_m - v_r}{v_0 + v_m - v_s} n. \tag{202}$$

240. Notes, tones, overtones, harmonics. Helmholtz has made the distinction that a *tone* (= German, *ton*) shall mean a sound not capable of further resolution into harmonic components; whereas a **note** is in general composed of a number of tones.

The pitch of a note is taken as that of its lowest tone, or fundamental. The fundamental is produced when the source vibrates as a whole. Thus a string having a node at each end and an antinode at its center would emit its fundamental tone (see Fig. 206a).

It is difficult to secure a perfectly pure tone. Most bodies vibrate in segments at the same time that they vibrate as a whole, and therefore produce a note compounded of its fundamental and of the partial tones, or **overtones,** due to the segments. Figure 206b shows a string emitting its fundamental and second harmonic. Overtones whose frequencies are integral multiples of the frequency of the fundamental are called **harmonics;** all others are called simply overtones. Thus, a rod free at both ends (Fig. 203) vibrates in such a way that if the frequency of its fundamental is taken as 1, those of the overtones are 2.92, 4.87, 6.32, 7.48, etc.* Here it is seen that the first and second overtones are approximately the third and fifth harmonics, but the others are not harmonics.

Harmonics and the corresponding terms of Fourier's theorem take their names from the fact that the fundamental and all harmonics up to the sixth, when sounding together, produce a harmonious note. The seventh and ninth harmonics, however,

* Watson, *op. cit.*, p. 390.

when sounded with the others, give unpleasant beats. Hence, in the piano the hammer strikes the string at a point approximately 1/8 of its length from the end, so as to produce an antinode at this point. The seventh and the ninth harmonics, which would have nodes at about this point, are made so weak as not to be disagreeable.

The presence of overtones is readily shown by resonance. Let a string be tuned to the frequency of one of the harmonics of another string. Then, when the latter is sounded, the former will absorb energy from it and will be set into **sympathetic vibration** with the harmonic to which it is tuned.

By means of a series of resonators, each of which may respond to a single frequency or may be adjustable throughout a short range of frequencies, the overtones of a given note may be isolated one at a time. The tone selected may then be amplified electrically to any desired intensity.

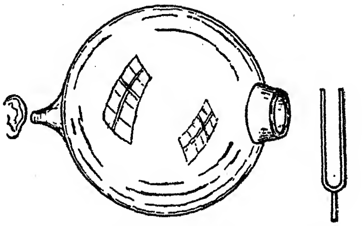


Fig. 202. Helmholtz Resonator

Resonators may be of various forms: rectangular, conical, cylindrical, etc. The spherical resonator designed by Helmholtz, shown in Fig. 202, is one of the best types.

241. Tuning forks. While any body vibrating with a proper frequency is a source of sound, the tuning fork, the stretched string, and the organ pipe are the sources usually studied first, because of their relative simplicity.

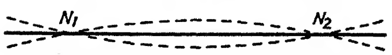


Fig. 203. Nodes in Vibrating Bar

The tuning fork consists of a bar, or rod, bent into the form of a long narrow U having a stem, or handle, at the base. If the bar before bending is supported so that it may vibrate freely, it will be found, when struck at an end, to vibrate with nodes at N_1 and N_2, as is shown in Fig. 203.

As it is bent toward the U-shape, these nodes approach the center and are finally located approximately as shown in Fig. 204. The point at which the stem is fastened continues, therefore, to be an antinode; hence the stem is raised as the prongs go outward

and lowered as they go inward. If the fork is mounted on a box made of thin wood and properly proportioned (see Fig. 194), this motion of the stem is communicated to the box, which is thereby thrown into sympathetic vibration and the sound is greatly augmented by resonance. Such boxes are called **resonance boxes.**

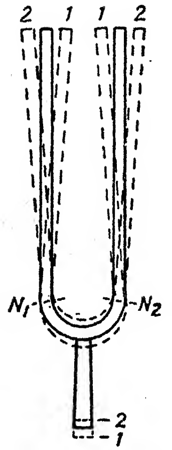


FIG. 204. Nodes in Tuning Fork When Sounding Fundamental

When a tuning fork is first struck, it vibrates in segments (Fig. 205) and emits overtones of high pitch, but these quickly die out almost completely. The note which then persists for a considerable time is very nearly the pure fundamental tone of the fork.

Tuning forks are easily maintained in continuous vibration by electrical means (sustained fork); the pitch varies but slightly with temperature; and the overtones are for most purposes negligible. Consequently, tuning forks are generally employed as **standards of pitch.** They are used also to regulate clocks, motors, and other devices for which great constancy of speed is desired.

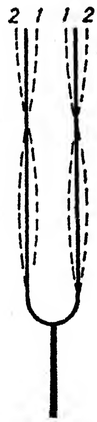


FIG. 205. Nodes in Tuning Fork When Sounding Overtones

242. Strings. It was stated in Sec. 224 that stationary waves occur whenever there is a sustained tone. In the case of a stretched string, the points at which the string is fastened are necessarily nodes, for they cannot move. The distance between nodes is a half wave length (Fig. 182). Hence the length L of a string is related to wave length λ by

$$L = \frac{\lambda}{2}. \tag{203}$$

Also, from Eq. (182),

$$v = n\lambda$$

and from Eq. (183),

$$v = \sqrt{\frac{F}{M}}.$$

Therefore,

$$n\lambda = \sqrt{\frac{F}{M}}$$

so that

$$n = \frac{1}{\lambda}\sqrt{\frac{F}{M}} = \frac{1}{2L}\sqrt{\frac{F}{M}}. \tag{204}$$

This *law of vibration of strings* may be stated as follows: The frequency of vibration of a stretched string varies directly as the

(a) (b)

FIG. 206. String Executing Fundamental and First Overtone

square root of the tension in the string and inversely as the length of the string and the square root of its mass per unit of length.

The law is readily verified by means of a **sonometer** (Fig. 207), with which each of the factors in Eq. (204) may be varied independently.

Besides vibrating as a whole (Fig. 206a), a string vibrates at the same time in segments, thereby producing overtones, which

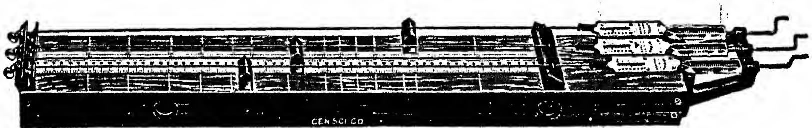


FIG. 207. Sonometer. (Courtesy Central Scientific Co.)

are very nearly harmonics. Figure 206b shows a string vibrating as a whole and also in two segments. Photographs of vibrating strings reveal that they actually vibrate in many segments.* This is shown also by the wave forms from the phonodeik. In these (Fig. 198), the irregularities as compared with a pure sine wave are due to the overtones.

In general, the richness of musical notes depends upon the abundance of harmonious overtones. The violin owes its preeminence among stringed instruments to the fact that it gives the musician very extensive control of the overtones.

243. Organ pipes. In a pipe organ, the vibrating body is the air within the pipes. The vibration of the material (wood or metal)

* Miller, *Science of Musical Sounds*, *op. cit.*, p. 66.

of which the pipes are made is so slight as to be negligible in an elementary discussion.

The construction of an organ pipe is shown in Fig. 208. The upper end may be open (**open pipe**), or it may be closed by an adjustable piston (**stopped pipe**).

With the stopped pipe the action is somewhat as follows. When air is blown in through the narrow slit M, it is deflected by the sharp edge E either into or out of the pipe. If it is deflected inward, additional air is forced into the column C, sending a compression up the pipe. This compression is reflected at the stopped end, returns down the tube to the large opening at O, and deflects the jet of air outward.

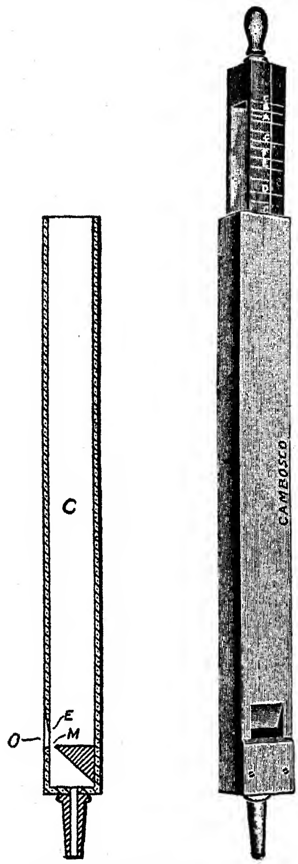


FIG. 208. An Organ Pipe

The effect of this current of air past the orifice O is then to produce a rarefaction in C, on the principle of the jet pump (Sec. 121). This rarefaction moves to the top of the pipe and is reflected; and on returning to the lower end of the pipe, it causes air to be drawn in so that a new compression starts upward.

The time for a complete wave (i.e., one compression and one rarefaction) to issue from the tube is the same as the time required for the disturbance to make two trips up and down the pipe. That is, the distance traversed by the wave front in one period is $4L$, where L is the length of the pipe; and by Sec. 215 this is the wave length. Hence, with the stopped pipe,

$$\lambda = 4L$$

which is the same result as was obtained with the resonance tube in Sec. 231.

If an open pipe is used instead of a closed pipe, the explanation is similar to the above, but it is somewhat complicated by the

fact that the compression and rarefaction are reflected at the upper open end with a change of 180° in phase.

The upper end of a stopped pipe is necessarily a node, and that of an open pipe, an antinode; but in both the opening at the lower end is large enough to make an antinode. By using the conventional means of representing compressional waves (Sec. 214), it is seen from Fig. 209 that a stopped pipe can produce only the harmonics whose frequencies have the ratios 1, 3, 5, etc.; whereas an open pipe can yield all the harmonics, 1, 2, 3, 4, 5, etc.

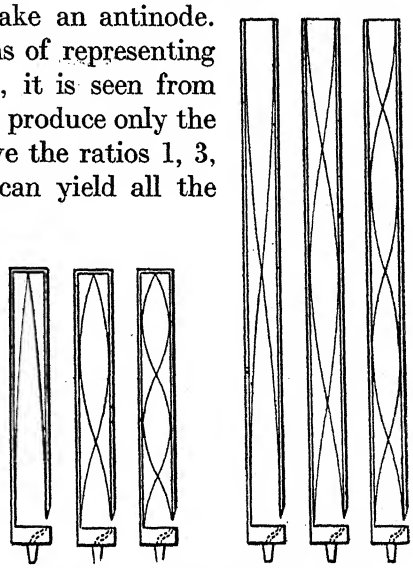


FIG. 209. Closed and Open Pipes

It will also be noticed that the length of a stopped pipe is 1/4 the wave length of the fundamental (first harmonic) tone, whereas the length of the open pipe is 1/2 the wave length of the fundamental. Hence a stopped pipe will need to be but half the length of an open pipe to produce the same fundamental tone. But **the notes from the open pipe will be much richer,** because they contain **all** the agreeable harmonics, whereas the stopped pipe contains only the **odd** ones, the 7th and 9th of which are disagreeable.

244. Musical scales. A study of musical tones reveals the fact that if their frequencies have the simple ratios, or **intervals,** 2 to 1, 3 to 2, 4 to 3, 5 to 3, 5 to 4, and 6 to 5, they will be pleasing when sounded together; that is, neither they nor their harmonics will produce disagreeable beats.

Tones whose frequencies have the ratios 4:5:6:8 are said to form a **major chord;** and those having the ratios 10:12:15:20, a **minor chord.**

A **major diatonic scale** is formed by taking any note of a major chord, building up new major chords above and below it, and selecting from the frequencies so obtained those that have approximately equal intervals. The note upon which the scale is built is called the **keynote,** or tonic. Minor scales are built up in a similar way.

While scales may be constructed upon any keynote, the so-called "natural scale" begins with middle C, to which physicists assign the frequency 256 because that is a power of 2, which simplifies computation. Musicians' concert pitch is based on a middle C of 264 vps. Below is given the major scale of C.

Major Diatonic Scale

Note	C	D	E	F	G	A	B	C′	D′
Frequency	264	297	330	352	396	440	495	528	594
	4		5		6			8	
Ratios				4		5		6	
		(3)			4		5		6

It will be noted that the frequency of C′ is twice that of C; i.e., the interval is 2 to 1. This fundamental interval is called an *octave*, because it comprises eight notes of a major diatonic scale.

In order to render selections in any desired key, it would be necessary to have scales from many keynotes. For even a few, this requires a very large number of tones—far more than could be produced on a piano or organ, which has a separate string or pipe for each tone.

To meet this difficulty, Johann Sebastian Bach introduced the **equally tempered scale.** In this, the octave is divided into 12 equal intervals, each of which is therefore the twelfth root of 2, or 1.0595 to 1. The equally tempered scale is now in universal use and can be distinguished from the natural scale only by those who have exceptional tone sensibility.

In Fig. 210 is shown the ordinary piano keyboard of 88 keys, with the staff and the musical symbol (note) for each tone.

245. The phonograph. The phonograph, invented in 1878 by Thomas A. Edison, was the first mechanical device for reproducing sounds in general, such as the human voice or the music of an orchestra. It consists essentially of a thin diaphragm, similar to that of a telephone receiver, closing the end of a horn which collects the sound. A stylus, or needle, is connected to the center of the diaphragm and presses against a cylinder of tin foil or some waxy compound. The diaphragm and stylus are moved along parallel to the cylinder by a screw, and as sound waves cause the diaphragm to vibrate, the needle traces a helical groove of varying depth in the waxen cylinder.

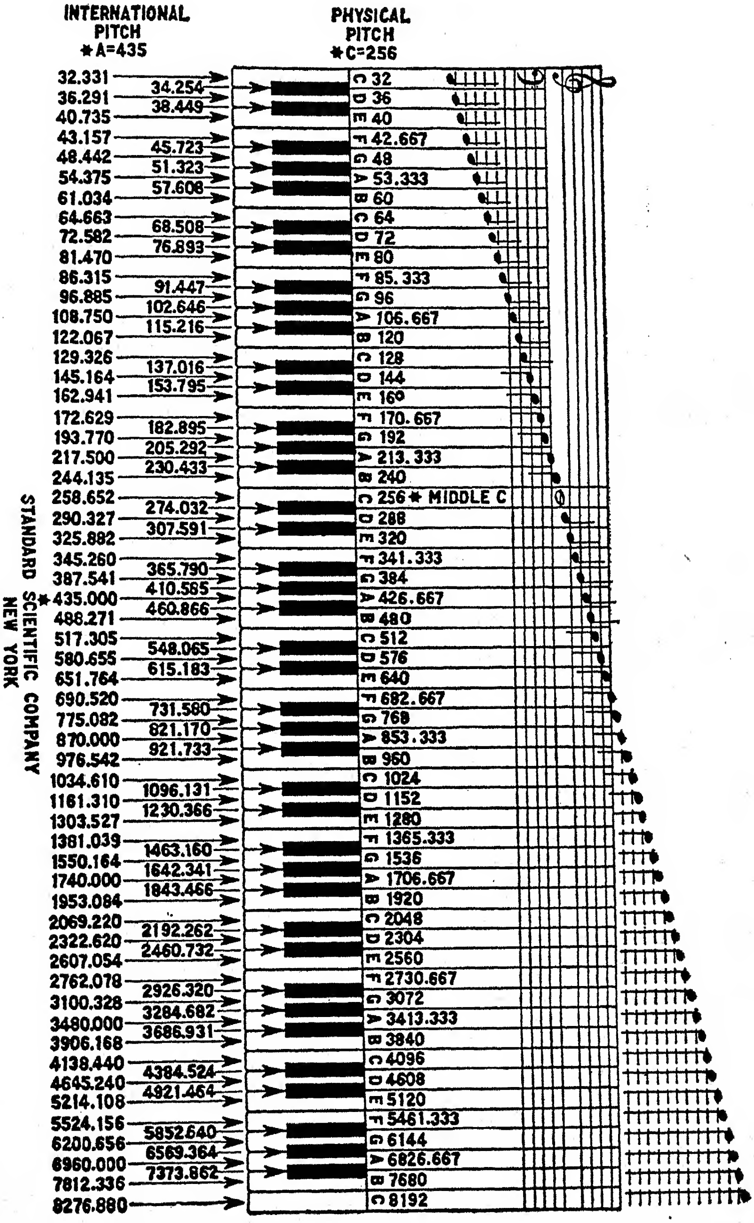


Fig. 210. Piano Keyboard. (Courtesy Chicago Apparatus Co.)

If the needle is then caused to retrace the groove, the recorded sound is reproduced quite accurately.

The **gramophone** of Emile Berliner employs the same basic principles, but the record is traced as a spiral groove on a flat circular plate rotating in a horizontal plane, the vibrations of the needle being sidewise instead of up and down.

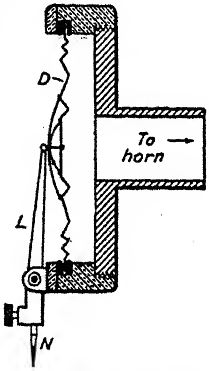


FIG. 211. Phonograph Reproducer

A reproducer of this type is shown in cross section in Fig. 211. *D* is a thin, circular, metal diaphragm corrugated in a special way to improve the reproduction of low notes. It is held around its outer edge between rubber washers. The vibrations are communicated to *D* from the needle *N* by means of the lever *L* and several light distributors at the upper end of *L*.

246. The ear is the primary sound-detecting instrument. Sound waves produce variations of pressure on the ear drum, or tympanic membrane (Fig. 212), causing vibrations of the drum. These vibrations are transmitted by a chain of three little bones (hammer, anvil, and stirrup) in the middle ear to a liquid contained in the cochlea, a chamber of snail-shell shape in the inner ear.

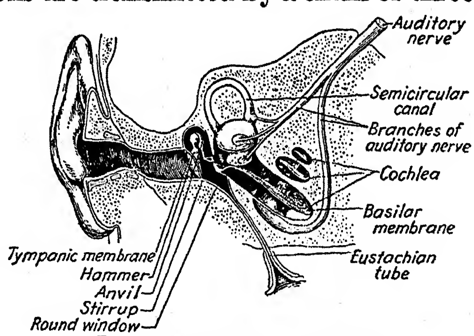


FIG. 212. The Ear. (Courtesy Henry Holt and Co.)

The cochlea is divided into two compartments throughout its length by the basilar membrane. This membrane is of fibrous structure, the fibers running crosswise from one side of the cochlea tube to the other. The fibrils of the **auditory nerve** are attached to the basilar membrane along its inner edge, where it joins the wall of the cochlea.

A vibration is communicated from the middle ear through an oval and a round window to the liquid on the two sides of the basilar membrane; and this latter experiences variations of stresses, thereby exciting the attached nerve fibrils, which transmit the

signal to the brain. This stimulus is thought to consist of an electric impulse, but the process is not thoroughly understood.

Notwithstanding its limitations, the performance of the ear is most remarkable. It can interpret as sound a vibration whose amplitude is approximately the diameter of an air particle (13×10^{-8}cm); and it can distinguish differences of pitch corresponding to less than 0.3% difference in frequency. The ear will register the reception of energy at the rate of 10^{-12} watt. It is therefore a most sensitive wattmeter. If energy were passed into a cubic centimeter of water at this rate and converted into heat without loss, it would take 130,000 years to raise the temperature 1°C.

The most astonishing property of the ear, however, is its ability to analyze complicated sounds. When the music of a band is heard, for example, the disturbance reaching the ear drum is the compressional wave due to all the instruments—merely changes of pressure. **But the hearing mechanism analyzes this disturbance, not simply into harmonic components as a harmonic analyzer does, but into the original complex notes of the individual instruments:** the part played by each instrument can be clearly distinguished in the infinitude of tones of the ensemble. How much of this resolution takes place in the ear, and how much in the brain, is a question in the field of psychology.

247. Audibility. Sounds that are just audible are said to be at

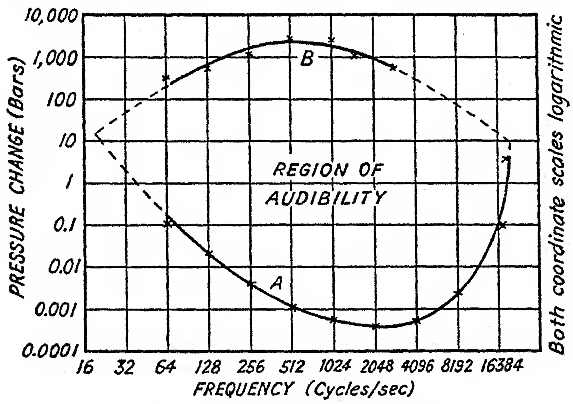


Fig. 213. Audibility Curves

the **threshold of audibility.** As the intensity of the sound waves (i.e., the energy transmitted per cm²/sec) is increased more and

more, the loudness increases until hearing ceases and the sensation becomes painful. The sound is then said to have reached the **threshold of feeling.** Between these two extremes lie all the vibrations that are perceived as sound.

Figure 213 shows the curves representing the threshold of audibility A and the threshold of feeling B for normal ears (72 persons). From these it will be seen that the frequency range within which sounds are audible is from about 20 to 20,000 complete vibrations (cycles) per second.

The ear is most sensitive (i.e., requires least power for an audible vibration) at about 2000 cycles per second; and at this frequency pressure changes * from .0006 to 2000 bars are audible.

248. Measurement of loudness: The decibel. If two sound waves, one having twice the intensity of the other, impinge successively upon the ear, the sensation produced by the first will *not* seem twice as loud as that produced by the second. That is, the loudness of sound as perceived by the ear is not directly proportional to the intensity of the stimulus, which in this case is the wave.

Loudness as judged by a large number of normal ears is known to depend upon both the intensity and the frequency, but there is no simple relation connecting them. Psychologists find that in general sensations obey approximately the **Weber-Fechner law: Sensations are proportional to the logarithms of the stimuli that produce them.**

Expressing this algebraically for sensations of sound,

$$S = c \log I + b \tag{205}$$

where S is the loudness of the sound sensation, approximately;
I is the intensity of the sound wave
c and b are constants depending upon the choice of units and axes.

* Sound intensity and pressure change are connected by the relation,

$$I = \frac{p_2}{415} \text{ at } 20°\text{C}.$$

where I is the intensity in microwatts per cm² and p is the pressure change in dynes per cm².

From Harvey Fletcher, *Speech and Hearing* (New York, D. Van Nostrand Company, 1929), p. 67.

Let S_0 and S be the loudness of two sound sensations produced by two sound waves of intensity I_0 and I, respectively.

Then by the Weber-Fechner law,

$$S_0 = c \log I_0 + b \text{ for the first sound.}$$
$$S = c \log I + b \text{ for the second sound.}$$

Since we really measure differences of loudness, we have on subtracting:

$$\alpha \equiv S - S_0 = c(\log I - \log I_0)$$
$$\equiv c \log \frac{I}{I_0} \qquad \text{(a)}$$

Again, for convenience, we choose our units so as to make $c = 1$.

In America, logarithms to the base 10 are chosen; then, by inspection, it is seen that if $\alpha = 1$, when $I = 10I_0$, c will $= 1$.

Thus,

$$1_{\text{bel}} = c \log_{10} \frac{10I_0}{I_0} \quad \text{makes } c = 1. \qquad \text{(b)}$$

This unit of α is named the **bel** after Alexander Graham Bell, the inventor of the telephone. Since $c = 1$ when these units are used, we may write

$$\alpha_{\text{bels}} \equiv \log_{10} \frac{I}{I_0}. \qquad (206)$$

α is called the **sound level** of the sound whose intensity is I with reference to the sound whose intensity is I_0.

From equation (b), it is seen that **the bel is the difference in the sound levels of two sounds when the intensity of one is 10 times the intensity of the other.**

The bel is too large for most practical purposes, hence the **decibel** (db) is defined:

$$1 \text{ decibel} \equiv 1/10 \text{ bel}.$$

Using decibels instead of bels, Eq. (206) becomes:

$$\alpha_{\text{db}} \equiv 10 \log_{10} \frac{I}{I_0}. \qquad (207)$$

It so happens that a change of 1 decibel in sound level gives about the smallest change in the sensation of hearing (loudness)

that the ear can detect. When used in this sense the decibel is sometimes called the "loudness unit," the "sensation unit," or the "phonic unit." But it should be borne in mind that this physical measure, sound level, does not represent accurately the magnitude of the sensation as perceived by the ear, just as the physical measure of temperature does not measure how hot a body feels to the hand.

Within the range from 700 to 4000 vps loudness is proportional to sound level quite accurately, but at lower frequencies loudness increases relatively faster than does sound level.*

The following table gives the sound levels of some familiar sounds in decibels, based on the threshold of hearing as reference level.

Sound	Sound Level (db)
(At threshold of hearing)	0
Cat purring at 3 ft	20
Whisper at 3 ft	30
Low speech	50
Ordinary speech	70
Loud speech	90
Pneumatic hammer at 5 ft	110

249. Reference levels of sound. In earlier work on loudness measurement, the reference level of sound intensity was taken as 1 microwatt per cm^2 ($= 10^{-6}$ watt/cm^2), but as this necessitated the use of negative values of sound level it has been generally abandoned. The preferred reference level is now taken as 10^{-16} watt/cm^2, which is about the threshold of audibility of the most sensitive ears.

It will be of interest to see how to change sound level measurements from one reference level to another.

Let I be the intensity of the waves producing the sound S, α be the sound level of S with reference to I_0 as zero level, and β be the sound level with reference to I_1 as zero level.

Then by Eq. (207),

$$\alpha = 10 \log_{10} \frac{I}{I_0}$$

$$\beta = 10 \log_{10} \frac{I}{I_1}.$$

* V. O. Knudsen, *Architectural Acoustics* (New York, John Wiley and Sons, 1932), p. 82.

Subtracting,

$$\alpha - \beta = 10\left(\log_{10}\frac{I}{I_0} - \log_{10}\frac{I}{I_1}\right)$$

$$= 10\log_{10}\frac{I_1}{I_0} = \text{Constant} \equiv \gamma$$

$$\alpha = \beta + \gamma.$$

Hence the sound level α of a certain sound with reference to a new zero level I_0, is equal to the sound level β of the same sound with reference to the old zero level I_1, plus a constant; and the constant to be added is the sound level of the old zero level with reference to the new zero level.

For example, in order to change sound level readings from the old reference level (10^{-6} watt/cm^2) to the new reference level (10^{-16} watt/cm^2), one must add to all the former readings the constant,

$$\gamma = 10\log_{10}\frac{10^{-6}}{10^{-16}}$$

$$= 100 \text{ db.}$$

250. Absorption coefficient. The reduction of noise is usually accomplished by the use of sound-absorbing materials. Smooth, hard surfaces in general reflect sound very efficiently; whereas rough, soft surfaces such as draperies, rugs, porous blocks, and cane fiber boards do not reflect well, but **absorb** sound very well. This is because the sound waves enter the pores and rugosities of the absorbing material and their energy is converted into heat by friction.

The **absorption coefficient, or absorbtivity** of a substance for sound is defined as the ratio of the sound energy that it absorbs to the total sound energy that falls upon it.

An open window absorbs all the sound that falls upon it, for all the sound goes out and none comes back. The absorption coefficient of an open window is therefore 1. Hence, the absorption coefficient of a substance equals the ratio of the amount of sound energy it absorbs to the amount absorbed by an open window of the same size. The absorption factors of some of the more common inside finishing materials are given in the following table.

ABSORPTION COEFFICIENTS FOR SOUND

Open Window	1.00
Brick	0.03
Carpet (heavily lined)	0.25
Concrete (unpainted)	0.02
Glass	0.027
Hard plaster	0.03
Heavy draperies (velour)	0.27
Wall board (cane fiber)	0.20
Wood (plain)	0.06
Wood (varnished)	0.03

The **absorption** (A_1) of an article is the product of its absorption coefficient a_1 by its area s_1.

It has been appropriately suggested * that the unit of absorption should be called the *sabin*, in honor of Professor W. C. Sabine of Harvard University, a pioneer in the science of architectural acoustics. **The sabin is the absorption for sound equal to that of an open window one square foot in area.** Absorption in sabins is therefore the absorption coefficient multiplied by the area in ft^2.

The absorption of chairs, seats, and auditors is frequently given in a table of absorption coefficients such as the one that follows. It should be understood that in such cases the value is per individual, and not per unit area.

ABSORPTIONS (IN SABINS)

Adult person	4.0 per person
Church seats (plain)	0.2 per seat
Church seats (cushioned)	1.0 to 2.0 per seat
Opera chairs	1.5 to 3.5 per chair

The **total absorption** (A) for a given room is the sum of the absorptions of its various interior surfaces and its furnishings. Thus,

$$A = A_1 + A_2 + A_3 + \cdots = a_1s_1 + a_2s_2 + a_3s_3 + \cdots \qquad (208)$$

where A will be in sabins if the areas are in ft^2.

251. Reverberation time. Reverberation is the prolongation of sound by repeated reflections. The rumbling of thunder is a familiar example of reverberation. From the practical point of

* V. O. Knudsen, *Architectural Acoustics* (New York, John Wiley & Sons, 1932), p. 127.

view, the most important feature about reverberation is its duration. This is called the reverberation time, or period.

The time of reverberation is defined as the time required for a sound to diminish from its initial intensity to one-millionth of that intensity; i.e., for the sound level to fall 60 decibels.

It has been shown * that for the most used frequency (512 vps) the time of reverberation is given by the relation:

$$t = 0.049\frac{V}{A} \tag{209}$$

where t is the time of reverberation in sec;
V is the volume of the enclosure in ft³; and
A is the total absorption in sabins.

252. Architectural acoustics. The term **acoustics** is often used as a synonym for the science of sound. As applied to buildings, it refers to the sound-reflecting and absorbing characteristics of halls and rooms. The acoustics of a room are said to be good if a speaker or music can be heard in a satisfying way throughout the room.

Prior to the year 1900, it was largely a matter of luck whether the acoustics of a room were good or bad. Chiefly through the researches of Professor Sabine, the factors affecting architectural acoustics are now well understood. Today an audience room may be designed with entire assurance that it will be suitable for its purpose.

From the preceding paragraphs, it will be clear that the acoustical efficiency of a room will depend upon its size and shape and upon the sound-absorbing properties of its contents. In general, if the period of reverberation is too long, the reflected sound interferes with the succeeding syllable coming direct from the source, as is shown in Fig. 214. Here the concave wall on the left concentrates the sound in a small area of the balcony, where it arrives slightly behind the direct wave from the speaker so that neither is intelligible.

If the reverberation time is too short, the absorption of the room is too great, and the room is said to be "dead." When it is correct, the reflected sound reinforces the direct sound, and a

* W. C. Sabine, *Collected Papers on Acoustics* (Cambridge, Harvard University Press, 1922), p. 43.

speaker or singer finds it easy to "fill" the room without voice strain.

From observations on a large number of auditoriums, it has been found that the best, or optimum, time of reverberation varies

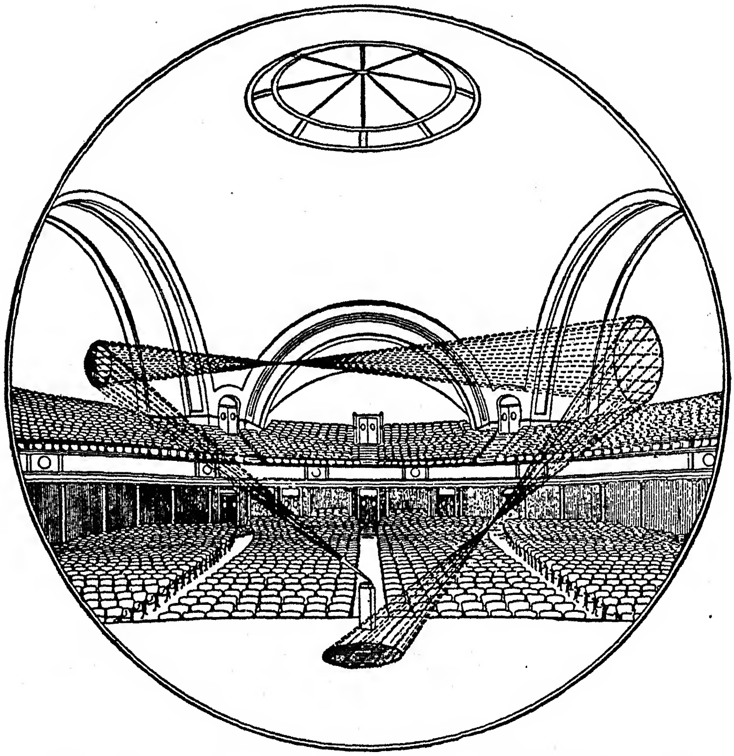


FIG. 214. Sound Foci in Auditorium. (Courtesy of Professor F. R. Watson, Univ. Illinois)

directly as the cube root of the volume of the room. From such data Professor F. R. Watson of Illinois University has prepared the curve of Fig. 215.

When a tentative design of an audience room has been made, its time of reverberation is calculated by means of Eq. (209) for the materials of construction contemplated. If this value is not quite close to the optimum value as given by the curve of Fig. 215, the interior finish and furnishings, and if necessary the design, are modified until the optimum is closely approximated.

With the use of modern public address systems employing mi-

crophones, vacuum tube amplifiers, and loud speakers properly placed, satisfactory acoustics may be obtained in almost any circumstances.

253. Supersonic vibrations. Longitudinal vibrations having frequencies of about 100,000 to 1,000,000 vps are far above the limit of audibility and hence are called *supersonic, or ultrasonic, vibrations.*

Supersonic waves may be obtained by an oscillating magnet, but they are usually produced by means of a piezoelectric crystal

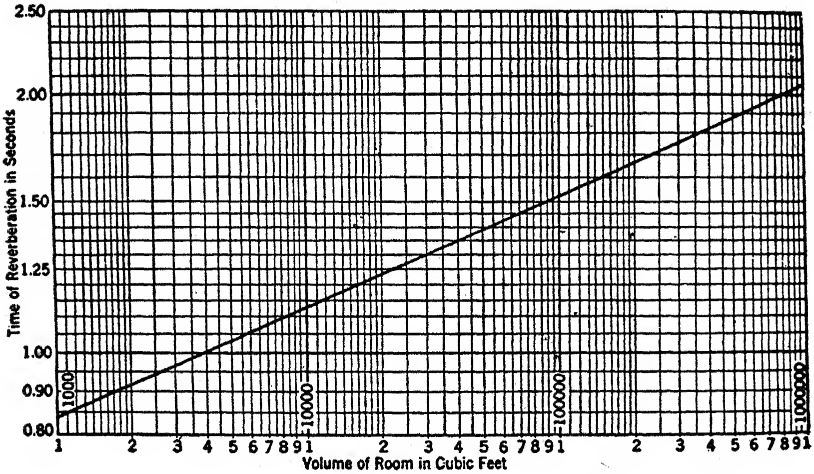


FIG. 215. Curve of Optimum Times. (Courtesy of Professor F. R. Watson, Univ. Illinois)

(quartz), mounted in a high-frequency oscillating circuit (Fig. 216). The circuit is tuned until the frequency of electric oscillations in the circuit containing the quartz plate is the same as the natural frequency of the plate for elastic vibrations. The crystal then executes vigorous mechanical vibrations at right angles to the direction of the electric field (see Sec. 370), and produces waves of the same frequency in the medium in contact with it on the end surfaces that are perpendicular to that vibration axis. A crystal 1 cm thick has a frequency of about 200,000 vps, and the frequency varies inversely as the thickness.

As was shown in Sec. 226, the energy transmitted per second by a wave varies as the square of the frequency. Supersonic waves

consequently transmit energy at a far higher rate than do sound waves.

On account of their great energies and frequencies, supersonic waves kill microscopic organisms and small aquatic animals, such as tadpoles. They also break up living cells and blood corpuscles in the liquids through which they are passed, and have been used to sterilize milk. If passed through globules of mercury in water, the mercury is dispersed in such fine particles throughout the water as to give it the appearance of black ink.

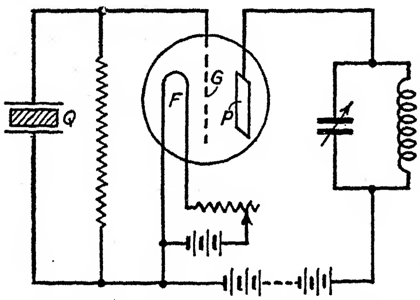


FIG. 216. Circuit for Supersonic Vibrations

By virtue of their short wave lengths (of the order of 1 mm), supersonic waves may be focused with mirrors of moderate size and have been used by ships for the detection of submarines and icebergs. A beam of the waves is sent out from the ship through the water; if it strikes an obstacle, part is reflected back to the ship. From the time required for it to make the round trip, the distance to the obstacle may be calculated.

PROBLEMS

1. Find the time required for sound to travel 3 mi in air at 70°F.
2. What is the wave length produced by a tuning fork having a frequency of 320 vps when the velocity of sound is 1100 ft/sec?
3. If the wave length produced in air by a tuning fork having a frequency of 256 vps is 4.36 ft, what is the velocity of the sound? What is the temperature?
4. A gun is fired, and 3 sec later the gunners hear the echo. If the temperature is 20°C, how far away is the reflecting surface?
5. If the temperature of the air is 25°C and a gun is fired in front of an abrupt cliff 1280 ft distant, how long after firing will the echo be heard by the gunner?
6. Three seconds after a stone is dropped into a well the sound of its splash is heard. If the temperature is 15°C, how deep is the well?
7. A man fires a pistol in a canyon and hears the echoes from the two canyon walls 1 sec and 2 sec, respectively, after the pistol was fired. If the temperature of the air is 18°C, what is the width of the canyon?
8. A wood chopper makes 12 strokes per minute. If the sound of each stroke reaches the observer as the axe makes the next stroke, and the temperature is 15°C, what is the distance of the chopper from the observer?

9. A wood chopper makes 15 strokes per minute. If the sound of each stroke reaches the observer as the axe makes the next stroke and the temperature is 20°C, what is the distance of the chopper from the observer?

10. A blow struck upon a cable was heard through the cable in 0.3 sec, and through the air in 1.2 sec. If the temperature was 15°C: (a) How far from the observer was the blow struck? (b) What was the speed of the sound in the cable?

11. A blow struck upon a cable was heard through the cable in 0.25 sec and through the air in 1.4 sec. The temperature was 20°C. (a) How far from the observer was the blow struck? (b) What was the speed of the sound in the cable?

12. Signals are sent from the shore to a ship by sound waves through the air and sea water, both at 10°C. If they reach the ship 0.6 sec apart, how far is the ship from the shore? (Velocity of sound in sea water is 1528 ft/sec.)

13. When a Kundt's tube is filled with air at 20°C, the distance between the centers of the dust heaps is 8.3 cm; and when filled with carbon dioxide, 6.4 cm. What is the speed of sound in carbon dioxide at 20°C?

14. If the dust heaps in a Kundt's tube filled with air are 20 cm apart when the temperature is 15°C, what are the wave length and the frequency? If the metal rod used is 60 cm long, what are the wave length and the velocity in the metal?

15. A glass tube placed with one end in a vessel of water is adjusted in length until it is in resonance with a tuning fork having a frequency of 256 vps. If the temperature is 25°C, what is the least length of the tube?

16. A glass tube, placed with one end in a vessel of water, is adjusted in length until it is in resonance with a tuning fork having a frequency of 128 vps. If the temperature is 15°C, what is the least length of the tube?

17. A tuning fork, when sounded with a fork having a frequency of 256 vps, gives 4 beats per sec. What are the possible frequencies of the unknown fork?

18. A tuning fork, when sounded with a fork having a frequency of 64 vps, gives 8 beats per sec. What are the possible frequencies of the unknown fork?

19. A tuning fork having a frequency of 400 vps is moved away from an observer with a velocity of 15 ft/sec. The temperature being 20°C, what is the frequency heard by the observer?

20. A tuning fork having a frequency of 420 vps is moved toward an observer with a velocity of 15 mph. The temperature being 30°C, what is the frequency heard by the observer?

21. Find the apparent pitch of a siren having a frequency of 800 vps when a car approaches with a speed of 40 mph. The wind is blowing from the observer in the direction of the car at a speed of 20 mph.

22. Find the apparent pitch of a train whistle having a frequency of 350 vps when the train approaches the station with a speed of 45 mph. The wind is blowing from the train toward the station at a speed of 15 mph.

23. One of the steel guy wires of an airplane is 8 ft long and .04 in.^{2} in cross section. If it emits a note of 128 vps when plucked, what is the tension in it?

24. What will the frequency of a wire 80 cm long and weighing 6 gm be when stretched with a force of 2 kg?

25. What must be the length of an organ pipe to give middle C, when the temperature is 15°C: (a) when the pipe is closed; (b) when the pipe is open?

26. What must be the length of an organ pipe to give middle C when the temperature is 20°C: (a) if the pipe is open; (b) if the pipe is stopped?

27. What must be the length of an organ pipe to give the octave above middle C, when the temperature is 25°C: (a) when the pipe is closed; (b) when the pipe is open?

28. Find the pitch (frequency) of the first two harmonics of a stopped pipe 8 ft long at 15°C.

29. A certain sound is produced by 5 times the power of a second sound. What is the difference of loudness in decibels? By what percentage must the sound power be increased in order to raise its loudness 1 db?

30. If average speech requires 10, loud speech 1000, and a soft whisper 0.001 microwatts, respectively, find their differences in decibels, using average speech as the zero reference level.

31. In the preceding problem, what would the intensity levels be if referred to a soft whisper as reference level?

32. An auditorium 60 ft by 50 ft by 20 ft is finished with plaster on lath, has a wooden floor, and contains 150 wooden seats. What is its reverberation period: (a) when the seats are empty; (b) when they are filled with auditors?

HEAT

The total energy of the universe is constant; its entropy tends to become a maximum.

—Clausius

CHAPTER XVIII

THERMOMETRY

254. The nature of heat. Up to the time of the American Revolution, heat was thought to be a weightless fluid called "caloric." This belief was upset in 1798 by Benjamin Thompson (later Count Rumford), a Tory of Woburn, Massachusetts. Having fled to Europe, Thompson, while boring brass cannon at Munich, succeeded in boiling water **without the aid of fire,** the heat being developed entirely by the friction of the boring tool against the brass. In 1799, Sir Humphry Davy melted ice by rubbing one piece against another in a vacuum. James Prescott Joule, in 1843, completed the overthrow of the caloric theory by showing that a definite amount of work produced a definite amount of heat.

Today we are sure beyond a peradventure that **heat is energy** of the molecules and atoms that compose a body; and we shall see presently that temperature is proportional to the average linear kinetic energy of the molecules for temperatures not too near absolute zero.

255. Temperature. Temperature is that condition of a body which determines its ability to transfer heat to other bodies. In everyday language, temperature is the measure of how hot (or how cold) a body is. When a silver spoon is placed in a cup of hot coffee, the spoon quickly becomes hot and the coffee cools a corresponding amount. In nature, heat always flows from the hotter to the colder body.

Temperature should be clearly distinguished from quantity of heat. If we have a bathtub full of water and dip out a cupful, the water in the cup has the **same temperature** as that in the tub; but there are probably a thousand cupfuls in the tub, and hence there is a **thousand times as much heat** in the tub as in the cup.

The sense of touch is not a trustworthy measure of temperature. If one hand is held in snow and the other in hot water for a few

minutes and both are then placed in a dish of tepid water, the former will tell us that the water is hot and the latter that it is cold. More dependable methods are easily devised.

When heat is added to a body its temperature, in general, rises. At the same time various other properties change, such as its length, its volume, its electrical resistance, its color, its thermo-emf with respect to another substance. The change in any of these properties may be used as a means of measuring the corresponding change of temperature; and all of them are actually so used.

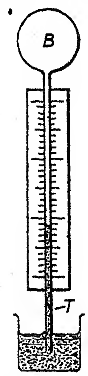


FIG. 217. Galileo's Thermometer

A thermometer is a device for measuring temperature. The metallic thermometer on a cooking stove employs change of length; the mercury thermometer, change of volume; the resistance thermometer, change of resistance; the thermocouple, change of thermo-emf.

The first thermometer was made by Galileo in 1593. As a thermometric substance he used air whose volume was the thermometric property. When the bulb B (Fig. 217) was heated, the air expanded, pushing downward the surface of the liquid in the tube T, the liquid being supported by atmospheric pressure and serving merely as an indicator of the volume of the air. The temperature was read on a scale placed behind the tube. This simple and sensitive arrangement has not survived as a thermometer because its reading varied with the barometric pressure as well as with the temperature.

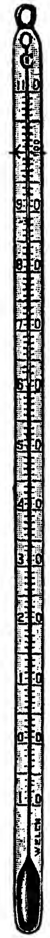


FIG. 218. Mercury-in-Glass Thermometer

Mercury as a thermometric substance was introduced by Boulliau in 1659, and it has many advantages. It is opaque; it does not stick to glass; it has a large coefficient of expansion; it is plentiful and easily purified.

The common form of mercury thermometer, shown in Fig. 218,

consists of a thin bulb blown on the end of a glass tube having a very small (capillary) bore and an outside diameter about that of a pencil. The tube is sealed off at the top while full of hot mercury. As the mercury contracts, it leaves a Torricellian vacuum above it; or else the space is filled with an inert gas if the temperature range is to be high.

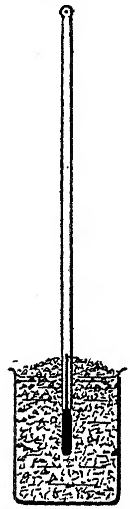


FIG. 219. Determination of Ice Point

After aging for a year or more, a thermometer is calibrated. First it is placed in crushed ice (Fig. 219) to a definite depth, and the point to which the mercury descends is marked on the tube. It is then placed with its bulb just above the boiling water in a "hypsometer," where it is surrounded by saturated steam at a pressure of 1 atmosphere (Fig. 220), and the point is marked to which the mercury rises.

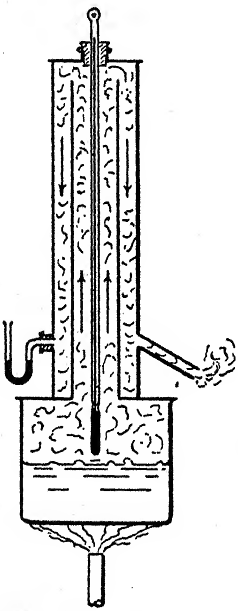


FIG. 220. Determination of Boiling Point

To make a centigrade thermometer, the lower point is marked 0° and the upper point 100°. The portion of the tube between them is then divided into 100 equal parts by means of a dividing engine. Celsius, who devised this scale in 1742, used these values but had them reversed.

To make a Fahrenheit thermometer, the lower point is marked 32° and the upper point 212°, and the intervening tube is divided into 180 equal parts.

The number chosen as the temperature of melting ice is called the **melting point of ice**; and that chosen as the temperature of the boiling water (or the steam over the water) is called the **boiling point of water at one standard atmosphere pressure.** Since these temperatures are accurately reproducible, they are called **fixed points.**

The foregoing is the general procedure in establishing any scale of temperature. It will be seen to consist of the following steps:

1. Choice of a thermometric substance; e.g., mercury.

2. Choice of a thermometric property; e.g., volume.
3. Choice of fixed points; e.g., melting point of ice and boiling point of water.
4. Choice of values for fixed points; e.g., 0° and 100°.
5. Choice of a relation between change of thermometric property and change of temperature.

The simplest relation that we could have between change of thermometric property and change of temperature is that of direct proportionality. Consequently, we choose as our **defining equation for temperature:**

$$\frac{t - t_1}{t_2 - t_1} = \frac{x - x_1}{x_2 - x_1} \tag{210}$$

where x_1 is the value of the thermometric property at the temperature t_1;
x_2 is the value of the thermometric property at the temperature t_2;
x is the value of the thermometric property at any temperature t;
(x_1, t_1) may be the lower fixed point;
(x_2, t_2) may be the upper fixed point; and
(x, t) is any point of the curve. (See Fig. 221.)

From analytic geometry, Eq. (210) is the standard equation of a straight line in terms of two points, (x_1, t_1) and (x_2, t_2). That is, in defining temperature **we assume a linear relation between temperature and the thermometric property.**

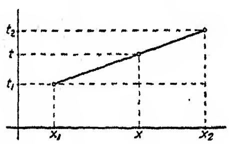


FIG. 221. Curve Defining Temperature

It should be noted that this physical measure of temperature does not measure how hot a body feels, say, to the hand. That is a matter of sensation and falls in the field of psychology.

256. Relation of centigrade and Fahrenheit scales. Eq. (210) gives us at once the relation between temperatures on the centigrade and Fahrenheit scales.

Let an uncalibrated thermometer (Fig. 222) be placed in melting ice and the length x_1 of the mercury column measured from any convenient datum plane. Let it then be placed in boiling water at one std. atms. pressure and the length x_2 of the mercury

column measured. Finally, let it be placed in a solution at any temperature t and the length of the mercury column measured from the same datum plane as for the two fixed points.

Now to make a **centigrade** thermometer, we assign the temperature values $t_1 = 0°$ and $t_2 = 100°$ to the fixed points, and call the temperature $t \equiv C$; so Eq. (210) becomes:

$$\frac{C - 0}{100 - 0} = \frac{x - x_1}{x_2 - x_1}. \qquad \text{(a)}$$

Then, to make a **Fahrenheit** thermometer, we assign the temperature values $t_1 = 32°$ and $t_2 = 212°$ to the fixed points, and call the temperature $t \equiv F$; so that Eq. (210) becomes:

$$\frac{F - 32}{212 - 32} = \frac{x - x_1}{x_2 - x_1}. \qquad \text{(b)}$$

But the right-hand members of Eqs. (a) and (b) are identical; hence,

$$\frac{C}{100} = \frac{F - 32}{180} \qquad (211)$$

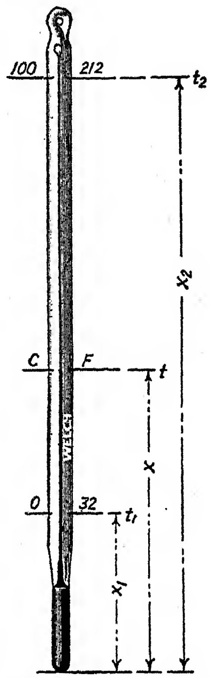


FIG. 222. Relation of Temperature Scales

which is the relation between the readings of a centigrade and a Fahrenheit thermometer for the same temperature.

257. Types of thermometers. Although the mercury-in-glass thermometer has the advantages mentioned in Sec. 255, which make it most convenient for measuring ordinary temperatures, it has serious disadvantages as well.

Mercury does not expand at a uniform rate: it expands more from 0° to 50° than from 50° to 100°C. The apparent change of volume is really the difference between the expansion of the mercury and that of the glass, and glass is a very uncertain material. The capillary bore is not uniform in diameter, and the bulb continues to shrink for years. The freezing point of mercury is −38.87°C, and its boiling point is 357.25°C, which limit its useful range—although, by introducing nitrogen above the mercury, its

upper limit may be extended to 600°C when made in glass, and to 750°C when made in quartz.

Alcohol may be used instead of mercury for temperatures down to −110°C, and pentane for those as low as −190°C. Using gallium instead of mercury, Elihu Thomson has constructed thermometers of quartz with a range up to 1000°C.

Special forms of thermometers are made for various purposes. **Six's maximum and minimum thermometer** is shown in Fig. 223. The bulb B is filled with alcohol, which is the thermometric substance. The lower part of the tube from A to C is filled with mercury, which serves simply as a flexible piston. Alcohol is also placed in the tube and bulb from C to D, but not enough to fill the space.

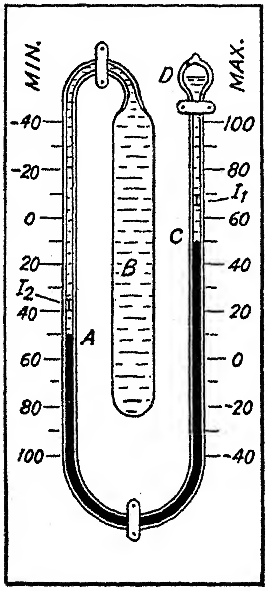


FIG. 223. Six's Maximum and Minimum Thermometer

As the temperature rises, the alcohol in B expands, pushing the mercury up the tube CD; the mercury in turn pushes ahead of it the small glass index I_1, which has an iron core. The lower end of this index shows at any time the maximum temperature.

When the temperature decreases, the alcohol in B contracts; the vapor in bulb D expands, pushing the mercury downward in CD, and the index I_2 is lifted by the mercury in A. The lower end of this index consequently shows the minimum temperature.

The indices are provided with delicate glass springs that hold them sufficiently to permit the alcohol to flow by, but not enough to resist the push of the mercury. They may be reset to the tops of the mercury columns by means of a small magnet which attracts their iron cores.

The **clinical thermometer** (Fig. 224) is another form of maximum thermometer. There is a constriction in its bore at C which permits the mercury to be forced upward, but which causes the mercury thread to break at this point when the temperature falls. The top of the mercury column therefore indicates the maximum

temperature reached. If the thermometer is swung quickly in a small circle, centrifugal force causes the mercury to pass downward through the constriction and reunite with that in the bulb. Care should be taken not to sterilize these thermometers in boiling water, as the scale seldom exceeds 110°F.

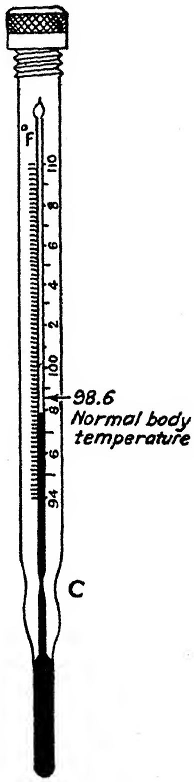


FIG. 224. Clinical Thermometer

The **bimetallic thermometer** (Fig. 225) takes its name from the fact that the spiral is made of two metals, usually brass on the inside and iron on the outside. The coefficient of expansion of brass is about 50% greater than that of iron. Hence, as the spiral heats up, it tends to straighten out and causes the pointer P to move over the scale, the end A being fixed to the case. This type of thermometer is frequently found on stoves and thermostats, where ruggedness is a prime requisite.

258. Standard thermometers. If the various thermometers mentioned in the preceding section are all carefully calibrated at 0°C and 100°C; if that interval is divided into 100 equal parts, as described in Sec. 255; and if they are then compared in the same liquid bath at various temperatures, say, −30°, 10°, and 50°C, as indicated by any one of them, they will be found *not in agreement.* Hence they are not suitable as standards.

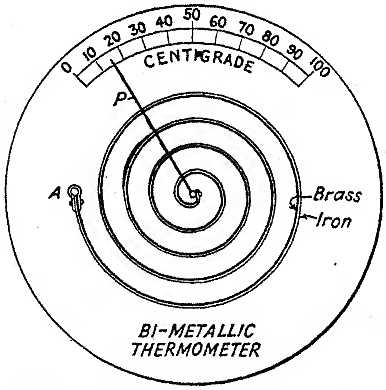


FIG. 225. Bimetallic Thermometer

Constant volume gas thermometers (Fig. 226), devised by Guillaume Amontons in 1702, when filled with any of the permanent gases, show excellent (though not exact) agreement. Pressure is the thermometric property employed in these thermometers, so

that its defining equation for temperature is:

$$\frac{t - t_1}{t_2 - t_1} = \frac{p - p_1}{p_2 - p_1}. \tag{212}$$

The constant volume gas thermometer is unwieldy and requires many refinements of manipulation. Hence it is not used except in standardizing laboratories.

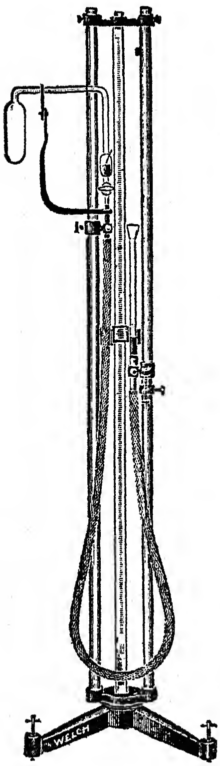


FIG. 226. Constant Volume Air Thermometer

The International Standard Thermometer is defined as a constant volume hydrogen thermometer. It must be made, filled, and operated according to standard specifications. Hydrogen was chosen as the thermometric substance because it does not react with the metals and has a very low boiling point. It serves from −250° to 500°C. Unfortunately, there is no material which will retain hydrogen above 500°. It would have been better to have chosen helium, which does not have this propensity and which has a boiling point lower than that of hydrogen—but when the selection was made, helium was not available in sufficient quantity. It is now used instead of hydrogen below −250°C.

From 500° to 1500°C the **constant volume nitrogen thermometer** is the standard; and at still higher temperatures, the **luminous filament optical pyrometer** (Sec. 314) is the standard.

The **platinum resistance thermometer** (see Sec. 411) is the recognized secondary standard from −195° to 650°C; and the **platinum platinum-rhodium thermocouple** (see Sec. 447), from 650° to 1063°C, the melting point of gold.

Though reasonably satisfactory as working standards, these thermometers are not in exact agreement, each yielding a scale of temperature characteristic of its thermometric substance.

Lord Kelvin, in 1848, devised a scale of temperature based on thermodynamic laws which is independent of any thermometric substance and therefore absolute. (See Sec. 328.) It may be shown

that it agrees exactly with the scale that would be given by a perfect gas.*

The absolute thermodynamic scale of temperature of Lord Kelvin is the standard scale of temperature.

There is no thermometer that gives temperatures on this scale by direct reading: they are determined by computation. However, their deviations from the readings of the constant volume gas thermometers are slight. (See International Critical Tables.)

PROBLEMS

1. What is the reading of a centigrade thermometer when a Fahrenheit thermometer reads 72°?

2. If a centigrade thermometer reads 20° in a certain solution, what would a Fahrenheit thermometer read in the same solution?

3. Normal body temperature is 98.6°F. What is it on a centigrade thermometer?

4. The highest recorded temperature from which the patient survived is 114°F. What is the equivalent temperature on the centigrade scale?

5. Ordinary solder, 2 parts lead and 1 part tin, melts at 240°C. What is its melting point in degrees Fahrenheit?

6. Wood's metal, an alloy of bismuth, lead, tin, and cadmium, melts at 150°F. What is its melting point in degrees centigrade?

7. If absolute zero is −273.14°C, what is it on the Fahrenheit scale?

8. The temperature of recalescence of iron is 778°C. What is it on the Fahrenheit scale?

9. At what temperature will the reading of a Fahrenheit thermometer be twice that of a centigrade thermometer?

10. At what temperature would a centigrade thermometer read 1/3 as much as a Fahrenheit thermometer?

11. In an air thermometer the pressure is 70.8 cm of mercury when in melting ice, and 96.8 cm when in boiling water at one atmosphere pressure. What is its temperature when the pressure is 85 cm of mercury?

12. The resistance of a platinum resistance thermometer is 12 ohms at 0°C and 16.4 ohms at 100°C. What is the temperature when its resistance is 18 ohms?

* Preston, *op. cit.*, p. 716.

CHAPTER XIX

EXPANSION

259. Expansion. If the molecules of all bodies are constantly in motion, as the kinetic theory of matter presumes (see Sec. 205), and if absolute temperature is proportional to the average linear kinetic energy of the molecules, then heating a body should in general increase the kinetic energy and hence the speed of the molecules; and it would seem probable that, the collisions being more vigorous, the average distance between molecules would increase. That is, when a body is heated we should expect it to expand, and most bodies do. Stretched indiarubber, iodide of silver, water between 0° and 4°C, and certain alloys are exceptions which are unexplained.

260. Coefficient of linear expansion. When a rod is heated uniformly along its entire length, each 1 cm of length will expand the same amount for 1° rise of temperature, for all the centimeters are alike. This amount is called the coefficient of linear expansion.

The coefficient of linear expansion is the change in length per unit of length at 0°C per degree change of temperature.

Let

l_0 be the length of the body at 0°C;
l_t be the length of the body at t°C;
t be the final temperature in degrees C; and
a be the mean temperature coefficient from 0° to t°C.

Then,

$l_t - l_0$ is the total change in length;
$\dfrac{l_t - l_0}{l_0}$ is change in length per unit length at 0°C; and
$t - 0$ is the total change in temperature.

Therefore, the change in length per unit of length at 0°C per degree change of temperature is

$$a \equiv \frac{l_t - l_0}{l_0(t - 0)} \tag{213}$$

and $$l_t = l_0(1 + at). \tag{214}$$

The coefficients of expansion of solids are generally so small and so nearly constant that for most purposes the computation of the final length may be based on the length l_1 at any temperature t_1 with ample accuracy.

Eqs. (213) and (214) then become

$$a \doteq \frac{l_2 - l_1}{l_1(t_2 - t_1)} \tag{215}$$

$$\doteq \frac{l_2 - l_1}{l_1 t}, \qquad \text{where } t \equiv t_2 - t_1$$

and $$l_2 \doteq l_1(1 + at). \tag{216}$$

It will be seen from Eq. (213) that the unit of temperature coefficient of expansion is

$$\frac{\text{cm}}{\text{cm} \times {}^\circ\text{C}} \quad \text{or} \quad \frac{\text{ft}}{\text{ft} \times {}^\circ\text{F}}$$

which are read "centimeters per centimeter per 1°C," or "feet per foot per 1°F." Here it will be seen that the units of length may be canceled out, and equally correctly we may say that the units of coefficient of expansion are:

$$\frac{1}{{}^\circ\text{C}} \quad \text{or} \quad \frac{1}{{}^\circ\text{F}}.$$

Since the units of length in the above cancel out, it follows that coefficients of expansion are not affected by the units in which the lengths are measured. But the units of temperature remain, consequently coefficients are different when different temperature units are used.

The temperature coefficient of expansion of a substance per 1°F is obviously 5/9 of its value for 1°C. Bearing this in mind, the above equations may be used as well with Fahrenheit as with centigrade temperatures.

COEFFICIENTS OF LINEAR EXPANSION PER 1°C *

Aluminum	25.5×10^{-6}
Brass	18.0
Glass (Pyrex)	3.0
Glass (soft)	8.5
Invar	0.9
Iron	11.9
Platinum	8.9
Quartz (fused)	0.54
Tungsten	3.4

The difference in the expansion of different metals for the same temperature change is well shown by the compound bar of

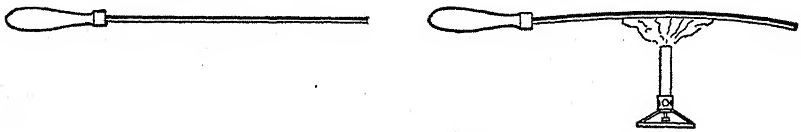


FIG. 227. Differential Expansion

Fig. 227. This consists of a thin strip of copper or brass welded to a thin strip of iron. As may be seen in the table, the coefficient of linear expansion for brass is greater than that for iron. Hence, when the bar is heated, it bends with the copper or brass on the outside of the curve. This was the principle employed in the bimetallic thermometer of Fig. 225. The thermostat (Fig. 228) likewise uses a bimetallic strip.

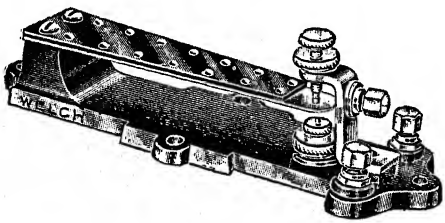


FIG. 228. Thermostatic Switch

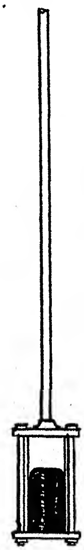


FIG. 229. Compensating Pendulum

In the same way, the balance wheel of watches is built of a bimetallic strip to compensate for temperature changes. Clock pendulums are frequently made compensating by means of a cylindrical bob containing mercury (Fig. 229). As the pendulum rod expands downward, the mercury expands upward just enough to maintain the center of mass unchanged in position.

* Mainly from Kaye and Laby's Tables.

On railroads, the expansion of metal bridges is provided for by having the trusses or beams rest on rollers at one end. Similarly, small gaps are left between the rails to allow for expansion.

261. Area and volume expansion. Quite analogous to the expression for the coefficient of linear expansion, we define the

$$\text{Mean coefficient of area expansion, } b \equiv \frac{A_t - A_0}{A_0 t} \tag{217}$$

and

$$\text{Mean coefficient of volume expansion, } c \equiv \frac{V_t - V_0}{V_0 t}. \tag{218}$$

From these, on clearing fractions, we have:

$$A_t = A_0(1 + bt) \tag{219}$$

and

$$V_t = V_0(1 + ct). \tag{220}$$

Here again the coefficients of expansion are characteristic of the substance and depend upon the scale of temperature (C or F), but they are independent of the units of area or volume, provided only that the same units are used on both sides of the equations.

Values of these coefficients are usually tabulated for 1°C. For Fahrenheit temperature their values are 5/9 of what they are for centigrade, since 1°F is 5/9 as large as 1°C.

The relations between the coefficients b and c and the linear coefficient a are easily calculated.

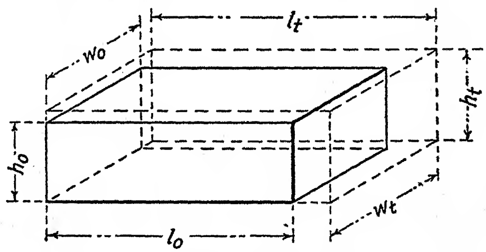


FIG. 230. Volumetric Expansion

Consider a body (Fig. 230) whose length, width, and height are l_0, w_0, and h_0, respectively, at the temperature 0°. Then at 0° its volume is: $V_0 = l_0 w_0 h_0$. On being heated to t°, each of these dimensions will expand, according to Eq. (214), and we shall have:

$$\begin{aligned} l_t &= l_0(1 + at) \\ w_t &= w_0(1 + at) \\ h_t &= h_0(1 + at) \end{aligned}$$

and the final volume V_t will be:

$$V_t = l_0 w_0 h_0 (1 + at)^3$$
$$= l_0 w_0 h_0 (1 + 3at + 3a^2t^2 + a^3t^3).$$

Since a is usually quite small, the terms $3a^2t^2$ and a^3t^3 are negligible compared to $3at$, except with very large values of t.

Hence,

$$V_t \doteq l_0 w_0 h_0 (1 + 3at)$$

The initial volume was

$$V_0 = l_0 w_0 h_0$$

and substituting these values in Eq. (220),

$$l_0 w_0 h_0 (1 + 3at) \doteq l_0 w_0 h_0 (1 + ct)$$

whence

$$c \doteq 3a. \tag{221}$$

That is, the volume coefficient of expansion is approximately equal to three times the linear coefficient for the same substance.

Similarly, it may be shown that the area coefficient of expansion is approximately equal to twice the linear coefficient, that is:

$$b \doteq 2a. \tag{222}$$

The **expansion of a plate having a hole in it** is of importance. The plate expands exactly as if the piece that originally filled the hole had not been taken out; and if this piece and the plate are heated to the same temperature, the removed piece will always fit the hole regardless of what that temperature may be.

FIG. 231. Ring Expands as if Solid. (Courtesy Cambridge Botanical Supply Co.)

This fact will be recognized at once when one recalls that the cast-iron piston of a steam engine fits the cast-iron cylinder just as well when hot as when cold. It may be nicely demonstrated by the ball and ring of Fig. 231.

At room temperature, the ball is a neat fit for the ring in any position and continues so, as long as the two are heated together. If the ring is heated alone, it becomes a loose fit; and if the ball is heated alone, it will not pass through the ring. The experiment

shows also that the material of the ball expands equally in all directions, which is characteristic of isotropic substances.

262. Anomalous expansion of water. Above 3.98°C, water, like most substances, expands with increase of temperature. But from 0° to 3.98°C it violates the general rule and contracts. Hence, **the specific volume,** i.e., the volume of unit mass, is a minimum at the temperature 3.98°C.

Density, being mass per unit volume, is the reciprocal of specific volume; hence, **density of water is a maximum at 3.98°C.** It will be recalled that this was the reason that the liter was defined at this temperature.

The variation of density is shown in Fig. 232.

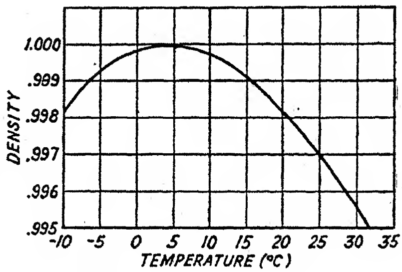


Fig. 232. Temperature-Density Curve for Water

This peculiarity in the behavior of water is a most fortunate provision of nature. As the water of a pond, for example, cools, the surface layer becomes denser and descends to the bottom until the whole body of water is at the temperature 3.98°C. On further cooling, the top layer becomes less dense and remains at the top until it freezes. Freezing therefore takes place from the top downward, leaving the water at the bottom at the temperature 3.98°C until the entire mass is frozen. Usually all the water does not freeze, and aquatic life can continue at this temperature.

If, however, water obeyed the usual law of expansion, rivers and lakes would freeze from the bottom upward. Many would freeze throughout every winter and might not thaw out completely during the summer. Aquatic life would to a large extent be destroyed. In temperate climates, skating would be safe but infrequent.

263. Absolute temperature. Our defining equation for temperature, Eq. (210), states that change of temperature is proportional to the corresponding change of the thermometric property. If pressure, for example, is the thermometric property, we obviously cannot change it to a smaller value after it has become zero. Hence we cannot measure a temperature lower than that cor-

responding to zero pressure, and this temperature is therefore the **absolute zero** on that scale.

The absolute zero on any scale of temperature is the temperature at which the thermometric property becomes zero.

Hence, by making the value of the thermometric property x in Eq. (210) equal to zero, we can calculate the temperature of absolute zero ${}_at_0$ on each scale. Writing ${}_at_0$ for t and 0 for x, we have:

$$\frac{{}_at_0 - t_1}{t_2 - t_1} = \frac{0 - x_1}{x_2 - x_1}$$

$${}_at_0 - t_1 = \frac{-x_1(t_2 - t_1)}{x_2 - x_1}$$

$${}_at_0 = -\frac{1}{\left(\dfrac{x_2 - x_1}{x_1(t_2 - t_1)}\right)} + t_1.$$

But by Eq. (213), $\dfrac{x_2 - x_1}{x_1(t_2 - t_1)}$ is the coefficient of expansion of the thermometric substance for the property which we are using. Call this coefficient a.

Then the temperature of absolute zero is:

$${}_at_0 = -\frac{1}{a} + t_1. \tag{223}$$

Hence, on any scale of temperature the absolute zero is the negative reciprocal of the coefficient of expansion for the thermometric property and substance used to establish that scale, plus the temperature of the lower fixed point. On the centigrade scale, the temperature of the lower fixed point being zero, the temperature of absolute zero is the negative reciprocal of the temperature coefficient, that is:

$${}_at_0 = -\frac{1}{a} \text{ for the centigrade scale.}$$

In the table below are given the temperature coefficients for the indicated property of various substances and the temperature of absolute zero on the corresponding centigrade scale.

Substance	Property	Temperature Coef.	Temp. Abs. Zero
Nickel	Resistance	0.0044	−227.3°C
Platinum	Resistance	0.00392	−255.1
Nitrogen	Pressure	0.003674	−272.18
Oxygen	Pressure	0.003672	−272.33
Hydrogen	Pressure	0.0036624	−273.04
Helium	Pressure	0.003662	−273.07
Mercury	Volume	0.000182	−5494.00
Iron	Length	0.0000119	−84000.00

From this table it will be seen that the permanent gases are in very good agreement, all giving absolute zero in the neighborhood of −273°C, and that the other types of thermometer are not in agreement. For this reason, and because of the fact that the zero as computed for the absolute thermodynamic scale of Lord Kelvin is −273.16 ± .01°C, we ordinarily use −273°C as the temperature of absolute zero. But it should be noted that this is **only an approximation.**

Absolute temperature on the Kelvin scale is therefore obtained by adding 273° to the temperature on the centigrade scale, and is indicated by the symbol *K*. Thus, at a pressure of 1 standard atmosphere, the temperature of boiling water is 100°C, or 373°K.

By Eq. (223) we find the temperature of absolute zero on the Fahrenheit scale to be −459.69°F. Consequently, **absolute temperature on the Fahrenheit scale** is obtained by adding 459.7° to the temperature Fahrenheit, and is indicated by the symbol *R*.* For example, the boiling point of water at 1 standard atmosphere pressure is 212°F or 671.7°R.

264. The law of Charles, or Gay-Lussac. While experimenting with balloons, Jacques Charles, a French physicist, observed in 1787 that the volume coefficient of expansion was about the same for all the ordinary gases. This fact was again discovered independently by Joseph Louis Gay-Lussac in 1802. About 1842, the very accurate experiments of Regnault gave the value of the coefficient as approximately 1/273.

* In honor of W. J. M. Rankine, Professor of Engineering in the University of Glasgow, Scotland.

Using this value in Eq. (220), we have, when P is constant

$$V_t = V_0(1 + ct)$$

$$= V_0\left(1 + \frac{1}{273}t\right) \tag{a}$$

$$= V_0\left(\frac{273 + t}{273}\right) \tag{b}$$

But $(273 + t)$ is our definition of absolute temperature. Calling

$$(273 + t) \equiv T$$
$$273 \equiv T_0$$
$$V_t \equiv V$$

and substituting in Eq. (b), we have:

$$V_t = \frac{V_0}{T_0}T. \tag{c}$$

Since V_0 and T_0 are both constants, we may call $V_0/T_0 \equiv C' =$ constant, so that

$$V = C'T$$

or

$$V \propto T \text{ when } P \text{ is constant.} \tag{224}$$

Eq. (224) is the **law of Charles, or Gay-Lussac: When the pressure of a given mass of a gas is constant, the volume is directly proportional to the absolute temperature.** Eq. (226) is also often called Charles' law.

265. The general law of an ideal gas. The laws of Boyle and Charles describe the behavior of an ideal, or perfect, gas when its mass is constant and its pressure, volume, and temperature vary two at a time. A general expression giving the value of any one of these properties when all three of the others vary may easily be deduced.

Let the pressure, volume, and absolute temperature of a given mass of a gas at a certain time be P_0, V_0, and T_0, respectively.

1. Keeping the pressure P_0 constant, let the absolute temperature be changed to T; then the new volume V_1 will be, by Charles' law:

$$V_1 = \frac{V_0}{T_0}T. \tag{a}$$

2. Keeping the temperature T constant, let the pressure be changed to P; then the new volume V will be given by Boyle's law:

$$PV = P_0V_1$$

$$= P_0\frac{V_0}{T_0}T$$

$$\frac{PV}{T} = \frac{P_0V_0}{T_0}. \qquad \text{(b)}$$

But $\frac{P_0V_0}{T_0}$ is a constant, which we shall call r, so that

$$PV = rT. \qquad \text{(c)}$$

3. If we keep P and T constant and double the mass M of the gas used in the foregoing discussion, we would obviously double V; hence r would have to be doubled in order that Eq. (c) may remain true.

Or, if we hold V and T constant and double the mass M of the gas used, we would obviously double P; hence, again r would have to be doubled to maintain the equation.

Thus r must contain M as a factor, and we may write:

$$r \equiv RM. \qquad \text{(d)}$$

Then the **general law of an ideal gas becomes:**

$$PV = RMT \qquad (225)$$

where R is the **gas constant for unit mass.**

The value of (R) depends upon the units chosen for the other quantities and, in the case of real gases, upon the kind of gas.

We may deduce a value of R as follows:

Let M be the mass of 1 gm-molecule (= 1 mole);
P_0 be the pressure of 1 std. atm. (= 1.0132×10^6 baryes)*; and
T_0 be the temperature of melting ice (= 273.16°K).

Then V_0 is the volume of 1 gm-molecule (= 22,415 cm³), and substituting these values in Eq. (225), we have:

$$R = \frac{1.0132 \times 10^6 \times 22415}{1 \times 273.16}$$

$$= 8.3143 \times 10^7 \text{ ergs per gram-molecule per } 1°K.$$

* Dynes per cm², commonly called "bars."

This value of R is a universal constant for all gases that closely approximate an ideal gas, provided that

M is expressed in gram-molecules, or moles;
P is expressed in baryes;
V is expressed in cubic centimeters; and
T is expressed in absolute degrees Kelvin.

If M is expressed in grams, the above value of R must be divided by the molecular mass * of the gas.

266. Other relations from the gas law. From the general law of an ideal gas, we can deduce two other useful relations.

From Eq. (b), Sec. 265,

$$PV = P_0V_0\frac{T}{T_0}.$$

If volume and mass are kept constant

$$V = V_0$$

and

$$\frac{P}{T} = \frac{P_0}{T_0} = \text{const.} \equiv C''. \qquad \text{(a)}$$

Then

$$P = C''T \qquad (226)$$

or

$$P \propto T \quad \text{when } V \text{ and } M \text{ are constant.}$$

This is often called Charles' law.

By definition,

$$T = t + 273 \text{ and } T_0 = 273.$$

Therefore

$$PV = P_0V_0\left(\frac{t + 273}{273}\right)$$

$$= P_0V_0\left(1 + \frac{1}{273}\,t\right)$$

$$= P_0V_0(1 + ct). \qquad (226a)$$

* Commonly called "molecular weight."

If we keep $V = V_0 =$ constant, and observe the variation of P with t, we have:

$$PV_0 = P_0V_0(1 + ct)$$
$$P = P_0(1 + ct). \qquad (227)$$

From Eq. (227) it is clear that for an ideal gas the temperature coefficient for pressure is the same as for volume; and experiment shows that this is very nearly true for actual gases. See tables.

267. The ideal gas. Free expansion. In the study of gases and the refined measurement of temperature, it becomes necessary to know whether the molecules of a gas exert forces of attraction or repulsion upon one another.

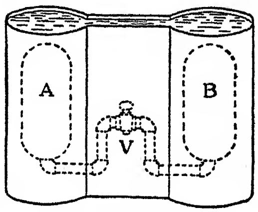


FIG. 233. Joule's Experiment

To determine this, Joule, in 1844, made the following **free expansion experiment.** Two copper vessels (Fig. 233), connected by a pipe, were submerged in a well-stirred water calorimeter. Vessel A contained air at 22 atmospheres pressure, and vessel B was evacuated. When valve V was opened, the air expanded into the vacuum of B until the pressure was the same in both vessels. During this free expansion (i.e., without external work), no change of temperature was observed.

However, the large quantity of water used tended to mask any slight temperature change. Consequently, in 1852, Joule and Thomson (Lord Kelvin) devised the more sensitive *porous plug experiment*, in which the gas expanded slowly through a plug of cotton wool from a region of high pressure to a region of low pressure. Air, oxygen, and carbon dioxide were cooled during the expansion; but at certain temperatures hydrogen was heated.*

From this experiment it appears that the molecules of actual gases exert a slight attractive force upon one another. In that case, work would be required to separate them; and this work, being done at the expense of the heat energy of the gas, would cause a slight cooling, different for different gases. Hence we define a perfect gas as follows:

* Edwin Edser, *Heat for Advanced Students* (New York, The Macmillan Company, 1936), p. 372.

An ideal, or perfect, gas would be a gas that would obey the general gas law and would exhibit no change of temperature on free expansion; i.e., its molecules would exert no forces of attraction or repulsion upon one another.

No such gas exists, but the so-called permanent gases (i.e., gases with very low critical temperatures) approximate perfect gases sufficiently closely for most practical purposes.

268. Van der Waals' equation. In general, real gases are more compressible than the gas law requires at ordinary temperatures and pressures, and are not compressible enough at high pressures; but for hydrogen PV is always slightly too high.

This is accounted for in part when we recall that in deducing Boyle's law (Sec. 205) by the kinetic theory of matter we did not take into consideration the diameter of the molecules or their mutual attractions for one another. This has been done to some extent in *Van der Waals' equation:*

$$\left(P + \frac{a}{V^2}\right)(V - b) = RT \tag{228}$$

where a and b are assumed to be constants characteristic of the gas. The term a/V^2 represents the pressure equivalent to the mutual attractions of the molecules; and b, the "co-volume," appears in theory to be about $4\sqrt{2}$ times the volume the gas would occupy if its molecules were in actual contact.

CONSTANTS OF VAN DER WAALS' EQUATION *

Gas	a (atmospheres)	b (liters per mole)
Hydrogen	0.19	0.00023
Oxygen	1.36	0.0316
Water	5.87	0.0332

Van der Waals' law expresses the behavior of real gases somewhat more accurately than does the general gas law, especially near their points of liquefaction.

* G. H. Cartledge, *Introductory Theoretical Chemistry* (Boston, Ginn and Company, 1929), p. 40.

269. Temperature and molecular kinetic energy. From Sec. 205 we have:

$$P = \tfrac{1}{3}nm\bar{v}^2$$

where n is the number of molecules in 1 cm^3;
m is the mass of 1 molecule; and
$\bar{v}^2$ is the average squared velocity of the molecules.

Multiplying both numerator and denominator by 2,

$$P = \tfrac{2}{3}n(\tfrac{1}{2}m\bar{v}^2). \tag{a}$$

But $(\tfrac{1}{2}m\bar{v}^2)$ is **the average kinetic energy of a molecule,** which we will call $\overline{E}$:

$$\overline{E} = \tfrac{1}{2}m\bar{v}^2 \tag{b}$$

so that

$$P = \tfrac{2}{3}n\overline{E}. \tag{c}$$

If we keep the original mass and volume of the gas constant and reduce the temperature from t°C ($= T$°K) to 0°C ($= T_0$°K), n would remain unchanged but P would decrease to P_0. Therefore $\overline{E}$ would decrease to a new value E_0, so that

$$P_0 = \tfrac{2}{3}n\overline{E}_0. \tag{d}$$

Dividing Eq. (c) by Eq. (d),

$$\frac{P}{P_0} = \frac{\overline{E}}{\overline{E}_0}. \tag{e}$$

But from Eq. (a) of Sec. 266,

$$\frac{P}{P_0} = \frac{T}{T_0}. \tag{f}$$

Therefore

$$\frac{T}{T_0} = \frac{\overline{E}}{\overline{E}_0}. \tag{229}$$

In words, **the absolute temperature of a given gas is proportional to the average kinetic energy of its molecules.***

From this law, it does not follow necessarily that the molecules of two *different* gases at the same temperature would have the

* It appears from statistical theory that this law is not true near the temperature of absolute zero.

same average kinetic energy. But if we consider Avogadro's law to be satisfactorily established by chemical evidence and the mass spectrograph (Sec. 508) and reverse the argument of Sec. 207, it is easily shown that

$$\frac{1}{2} m_1\bar{v}_1^2 = \frac{1}{2} m_2\bar{v}_2^2$$

for the same value of T, where m_1, m_2, and $\bar{v}_1^2$, $\bar{v}_2^2$ are the masses and average squared velocities, respectively, of two different gases.

This is *Maxwell's law* which he deduced independently from the kinetic theory of matter (Sec. 206): **At the same temperature, the molecules of all gases have the same average kinetic energy.***

PROBLEMS

1. Assuming the highest summer temperature to be 35°C and the lowest winter temperature to be −20°C, what allowance should be made for the expansion of a 1200-ft steel bridge?

2. Assuming the highest summer temperature to be 40°C and the lowest winter temperature to be −10°, what allowance should be made for the expansion of a 40-ft steel exposed beam?

3. A wheel is 2.5 m in circumference at 15°C. An iron tire measures 2.491 m around its inner circumference. To what temperature must the tire be raised in order that it may just slip on?

4. A steel tape was measured at 20°C and found to be 99.78 ft long. At 40°C its length was found to be 100.04 ft. What was its coefficient of expansion? At what temperature would it be exactly 100 ft long?

5. If the temperature coefficient of expansion of brass is 19×10^{-6} per 1°C, what is it per 1°F?

6. If steel rails are 30 ft long and have a coefficient of expansion of 11×10^{-6} per 1°C, what space must be left between the ends of rails to allow for expansion when the temperature varies from −10°F to 70°F?

7. Calculate the area at 70°F of a plate of sheet iron which is 6 ft long and 3 ft wide at 32°F.

8. A glass flask holds 1240 gm of mercury at 15°C. How much mercury will overflow when the whole is heated to 95°C?

9. At 68°F the volume of a balloon is 300 yd³. If no gas is allowed to escape, what will be its volume at a height at which the pressure is 0.75 atmosphere and the temperature 41°F?

10. A liter of dry air at standard conditions weighs 1.296 gm. What would be the mass of 3 liters at 115°C and a pressure of 4 atmospheres?

* It appears from statistical theory that this law is not true near the temperature of absolute zero.

11. What will be the mass of a cubic meter of air at 60°C and a pressure of 45 cm of mercury?

12. What would be the volume at standard conditions of a mass of gas which at 77°C and 1.1 atmospheres pressure occupies a volume of 12 liters?

13. Compute the value of the gas constant R for air when pressures are in grams per square centimeter, volumes in cubic centimeters, masses in grams, and temperatures in degrees centigrade.

14. Compute the value of the gas constant R when pressures are in pounds per square inch, masses in pounds, and temperatures in degrees Fahrenheit.

15. A liter of dry air at standard conditions weighs 1.296 gm. What would be the mass of 5 liters at 47°C and a pressure of 10 atmospheres?

16. What is the mass of 8 m³ of air at 60°C and a pressure of 58 cm of mercury?

17. An automobile tire has an average diameter of 24 in. and a cross-sectional diameter of 5 in. Supposing the tire empty, what volume of air measured at 1 atmosphere pressure must be added to inflate the tire to 34 lb gauge pressure?

18. What is the mass of atmospheric air in a room 9 ft by 14 ft by 16 ft when the barometer reads 77 cm of mercury and the temperature is 20°C?

19. An auditorium is 22 ft by 70 ft by 100 ft. What mass of air does it contain when the temperature is 15°C and the pressure 75 cm of mercury?

20. What is the draft in inches of water produced by a chimney 30 m high, if the average temperatures of the air outside and inside are 10°C and 85°C, respectively? ("Draft" is the difference of pressure outside and inside at the bottom of the chimney.)

21. The observed reading of a barometer at 20°C is 71 cm. What correction must be made to this to reduce it to 0°C, if the volume coefficient of expansion of mercury is 0.000182 and the linear coefficient of brass 0.0000184 per 1°C?

$\Delta L = L_0 \alpha \Delta T$

CHAPTER XX

CALORIMETRY

270. Heat units. The process of measuring quantity of heat is called **calorimetry.** Before we can proceed with calorimetry, it is necessary to define units of quantity of heat.

The **calorie** (also called the gram-calorie, the lesser, or the normal calorie) is the amount of heat required to raise the temperature of 1 gm of water from 14.5° to 15.5°C.* It is equivalent to 4.185 joules.

The **greater calorie** (or kilogram-calorie) is defined as 1000 calories. The "food calorie" used in expressing the energy values of different foods is always the greater calorie.

The **British thermal unit** (Btu) is the amount of heat required to raise the temperature of 1 lb of water from 59° to 60°F.

$$1 \text{ Btu} \equiv 1 \text{ lb} \times 1°\text{F} = 453.6 \text{ gm} \times \tfrac{5}{9}°\text{C} = 252 \text{ calories.} \tag{230}$$

271. Specific heat. The specific heat *c* of a substance is the number of calories required to raise the temperature of 1 gm of the substance 1°C, or the number of Btu required to raise the temperature of 1 lb of the substance 1°F.

Numerically this is the same in both systems of units. For if the specific heat of a substance is c(cal/gm°C), we have, on reducing each unit to its British equivalent:

$$c\frac{\text{cal}}{\text{gm°C}} = c\frac{1 \text{ cal}}{1 \text{ gm} \times 1°\text{C}} = c\frac{1/252 \text{ Btu}}{1/453.6 \text{ lb} \times 9/5°\text{F}} = c\frac{1 \text{ Btu}}{1 \text{ lb} \times 1°\text{F}}.$$

Thus the number of calories required to raise the temperature of 1 gm of a substance 1°C is the same as the number of Btu required to raise the temperature of 1 lb of the same substance 1°F.

The **specific heat of water** at 15°C is 1 cal/gm°C by the defini-

* Two other calories are sometimes used:

1. The *20° calorie* is the amount of heat required to raise 1 gm of water from 19.5° to 20.5°C (= 4.182 joules).

2. The *mean calorie* is 1/100 the amount of heat required to raise 1 gm of water from 0° to 100°C (= 4.1853 joules).

tion of the calorie. It was assumed to be unity at all temperatures until Henry A. Rowland, in his famous experiment on the mechanical equivalent of heat, proved that it is not. The variation

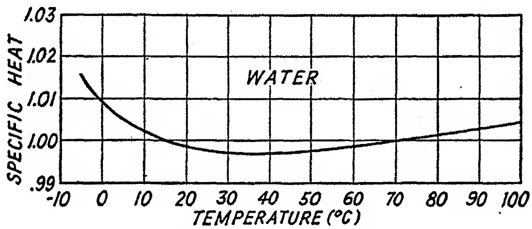


FIG. 234. Specific Heat of Water

of the specific heat of water with temperature based on the 15°-calorie is shown in Fig. 234.

TABLE OF MEAN SPECIFIC HEATS *

Substance	Temp. Range	Specific Heat
Aluminum	17–100°C	0.217
Copper	15–100	0.09305
Hydrogen (liq.)	−253	6.0
Ice	−21 to −1	0.502
Iron	18–100	0.113
Lead	20–100	0.0305
Lithium	0–100	1.09
Mercury	20–40	0.03318
Silver	15–100	0.0563
Steam (superheated)	100–500	0.48
Zinc	20–100	0.0938

272. Thermal capacity. Water equivalent. The *thermal capacity* of a body is the amount of heat required to raise the temperature of the body 1°C or 1°F, as the case may be. If the mass of a body is M and its specific heat, c, then

$$\text{Thermal capacity of body} \equiv cM \text{ cal or Btu.} \qquad (231)$$

Clearly the thermal capacity of a body will not be the same in the cgs and fps systems because the values of M will be different.

The **water equivalent** of a body is the mass of water that requires as much heat to raise its temperature one degree as is

* Chiefly from *Hand Book of Physics and Chemistry* (The Chemical Rubber Co.).

required to raise the temperature of the body one degree. From the definitions of the calorie and the Btu,

$$\text{Water equivalent of body} \equiv cM \text{ gm or lb mass} \qquad (232)$$

where c is a pure number numerically equal to the specific heat.

273. Quantity of heat. By the definition of the preceding section, the thermal capacity cM of a body is the amount of heat required to change its temperature 1°. Hence, if the temperature of a body rises from t_1 to t_2, the

$$\text{Quantity of heat gained} = cM(t_2 - t_1) \text{ heat units.*} \qquad (233)$$

If the temperature of the body falls, a similar expression will represent the amount of heat lost by the body.

274. The method of mixtures. One of the best methods of determining the specific heat of a solid or a liquid is the following, known as the *method of mixtures.*

274

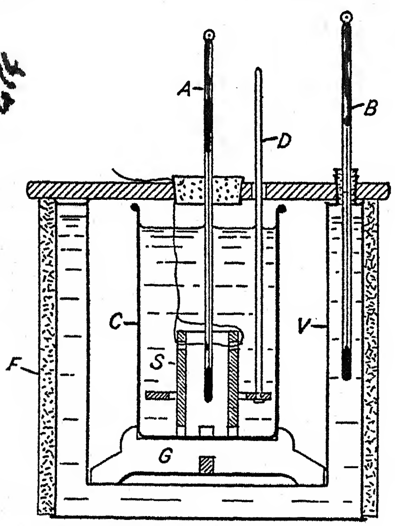


FIG. 235. Mixture Calorimeter

In Fig. 235, the inner vessel C of a calorimeter is nearly filled with a known mass of water at a known temperature, and is placed inside a constant-temperature enclosure, or outer vessel, V. This outer vessel is double-walled, filled between these walls with water, and lagged with a heavy layer of felt F. The specimen S is heated to a known temperature and lowered quickly into C without splashing. The mixture of the water, the specimen, the stirrer D, the thermometer A, and the inner vessel, thus come to a new steady temperature observed on A. The specific heat is then computed as follows:

Let M be the mass of the specimen;
M_1 be the mass of the inner vessel and the stirrer;
M_2 be the mass of the water in the inner vessel;

* Provided the body does not undergo a change of state. See Chap. XXI.

c be the specific heat of the substance of the specimen;
c_1 be the specific heat of the substance of the inner vessel and the stirrer;
c_2 be the specific heat of the water;
t be the initial temperature of the specimen;
t_1 be the initial temperature of the inner vessel, stirrer, water, and thermometer; and
t_2 be the final temperature of the inner vessel and its contents.

Then, cM is the thermal capacity of the specimen;
c_1M_1 is the thermal capacity of the inner vessel and the stirrer;
c_2M_2 is the thermal capacity of the water;
z is the thermal capacity of the thermometer (previously determined);
$cM(t - t_2)$ is the heat lost by the specimen;
$c_1M_1(t_2 - t_1)$ is the heat gained by the inner vessel and stirrer;
$c_2M_2(t_2 - t_1)$ is the heat gained by the water; and
$z(t_2 - t_1)$ is the heat gained by the thermometer.

If we assume that the specimen is transferred from the heater to the inner vessel so quickly that its loss of heat in transit is negligible, and if we believe in conservation of energy, we may write the equation:

$$\text{Heat lost} = \text{Heat gained}$$
$$cM(t - t_2) = (c_1M_1 + c_2M_2 + z)(t_2 - t_1)$$

whence

$$c = \frac{(c_1M_1 + c_2M_2 + z)(t_2 - t_1)}{M(t - t_2)}. \tag{234}$$

275. Calorific value. The calorific value of a fuel is defined as the number of heat units liberated by the complete combustion of unit quantity of the fuel.

With solid and liquid fuels, *unit mass* is used. The determination is made in a calorimeter similar to that of Fig. 235, but provided with a watertight bomb B (Fig. 236), in which the fuel is burned in an atmosphere of oxygen.

The quantity of a gas that could be got into the bomb is so small that tests of gases are usually made with a **continuous flow calorimeter,** and the quantity is **unit volume** at standard condi-

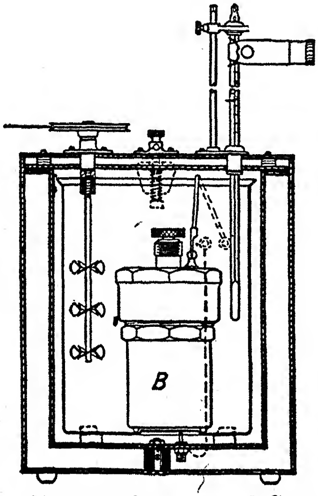


FIG. 236. Oxygen Bomb Calorimeter. (Courtesy Park Calorimeter Co.)

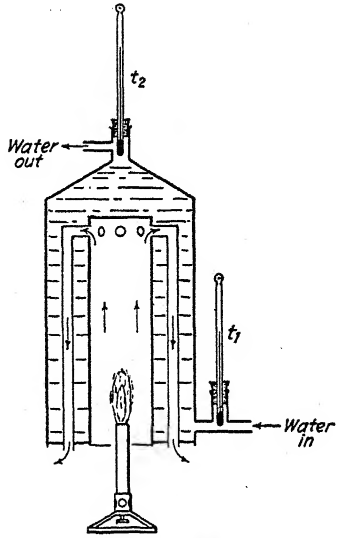


FIG. 237. Continuous Flow Calorimeter

tions. Figure 237 shows in cross section Mann's modification of Junker's continuous flow calorimeter.

The gas is burned continuously under the calorimeter, and the hot products of combustion pass outward as shown, through the tubes which are surrounded by water. This water flows through the calorimeter at a steady rate.

Let V be the volume * of gas burned after conditions have become steady;
M be the mass of water passing through the calorimeter after conditions have become steady;
t_1 be the initial temperature of water;
t_2 be the final temperature of water;
c be the specific heat of water; and
H be the calorific value of the gas.

Then the total heat received by the circulation water is $cM(t_2 - t_1)$.

* Reduced to standard conditions.

Since the calorific value H is defined as the heat liberated per unit volume, the

$$\text{Total heat liberated} = HV.$$

And if the average temperature of the water at inlet and outlet is the same as room temperature, radiation losses may be neglected so that

$$HV = cM(t_2 - t_1) \tag{235}$$

and

$$H = \frac{cM(t_2 - t_1)}{V} \text{ cal/cm}^3 \text{ or Btu/ft}^3.$$

Adjustment should be made so that the products of combustion leave the calorimeter at room temperature.

PROBLEMS

1. What is the thermal capacity of 10 lb of iron?

2. What is the water equivalent of a nickel vessel weighing 150 gm and a copper stirrer weighing 40 gm?

3. How much heat is required to raise the temperature of 65 gm of mercury from $-5°C$ to its boiling point at atmospheric pressure?

4. From the definitions, compute the relation between a calorie and a Btu.

5. Starting with the definitions, compute the number of Btu in 1 kg-cal.

6. The calorific value of a certain bituminous coal is 14,500 Btu per lb. Calculate the value in calories per gram.

7. A piece of copper at 100°C is dropped into 1500 gm of water and raises its temperature from 20°C to 22.5°C. What is the mass of the copper?

8. If the specific heat of glass is 0.199 and that of mercury is 0.033, compute the water equivalent of a thermometer when 2 gm of glass and 8 gm of mercury are submerged.

9. A copper kettle weighing 1 kg contains 5 kg of water at 20°C. Find the amount of heat necessary to raise the temperature of the water to the boiling point when the barometer reads 76 cm of mercury.

10. A teacup whose mass is 150 gm and temperature 15°C has 125 cm^3 of boiling water poured into it, and the final temperature is 82°C. What are the water equivalent and the specific heat of the teacup?

11. A metal ball weighing 90 gm is heated to 100°C and dropped into a copper calorimeter having a mass of 80 gm and containing 100 gm of water at 10°C. If the final temperature is 16°C, what is the specific heat of the metal?

12. A piece of silver, whose mass is 70 gm, is heated to the temperature of a furnace and then dropped into a copper calorimeter (mass 80 gm) containing 110 gm of water at 20°C. If the final temperature of the water is 35°C, what is the temperature of the furnace?

13. A metal ball weighing 80 gm is heated to 100°C and dropped into a copper calorimeter having a mass of 60 gm and containing 150 gm of water at 15°C. If the final temperature is 20°C, what is the specific heat of the metal?

14. How much heat will be required to raise the temperature of the air in a room 5 m by 5 m by 3 m from 0°C to 20°C, when the pressure is 1 std. atm.?

15. A sample of coal weighing 1.1 gm is burned in a steel bomb whose mass is 1500 gm. The bomb is submerged in 2000 gm of water in a copper calorimeter whose mass is 600 gm. If the rise in temperature is 3.4°C, what is the calorific value of the coal?

16. In a continuous flow calorimeter, the combustion of 1180 cm^3 of gas raises the temperature of 755 gm of water from 19°C to 38.5°C. Find the calorific value of the gas.

17. A specimen of gasoline weighing 0.1 oz is burned in a fuel calorimeter containing 7.5 lb of water. If the calorimeter is of copper weighing 3.6 lb and the temperature rises from 63°F to 78.7°F, find the calorific value of the gasoline in Btu per lb.

18. When 2.5 ft^3 of city gas were burned in a continuous flow calorimeter, 125 lb of water were raised from 61°F to 73°F. If 18 Btu were carried off in the products of combustion, what was the calorific value of the gas?

CHAPTER XXI

CHANGE OF STATE

276. Kinetic theory. We have seen that in accord with the kinetic theory of matter, a body expands when its temperature increases, justifying the belief that the mean distance between molecules increases. It might be expected, therefore, that if we continue to raise the temperature, the mean distance between them would become so great that a molecule would no longer be restrained to its original molecular neighborhood by the attractions of those neighbors, but might wander with many collisions through the entire mass, like bees in a swarm. Cohesion would be so greatly reduced that the body might not even retain its definite solid shape—and, as we know, it does not, but conforms to the shape of the containing vessel.

The phenomenon of changing from the solid state, or phase, to the liquid state is called **melting, or fusion.**

277. Melting point. If a beaker filled with crushed ice at, say, $-10°C$ is heated over a burner, it will be found that melting will begin when some of the ice reaches 0°C. Thereafter, the mixture of ice and water, if well stirred, will remain at this temperature until all the ice is melted.

The temperature at which a substance changes from the solid to the liquid state is called its melting point. Every pure, crystalline substance has a definite melting point; * but amorphous substances like glass, sealing wax, and butter have not.

278. Heat of fusion. In the above experiment of melting ice, after melting began, the temperature remained constant until all the ice was melted, notwithstanding the fact that we were adding heat all the time at a constant rate. All pure, crystalline substances behave in this way. The heat added during the process of

* Provided it does not decompose before melting.

melting appears to be necessary to effect the change of state only, and does not affect the temperature.

The amount of heat required to change a unit mass of a substance from the solid state to the liquid state without change of temperature is called its heat of fusion. It is characteristic of the substance.

Heats of Fusion
(in cal/gm)

Copper	43.0
Ice	80.0
Iron, cast	23.0
Lead	5.4
Mercury	2.8
Tin	14.6
Zinc	26.6

Since the heat of fusion does not raise the temperature of the substance, it seems to disappear; hence for a long time it was called the "latent (i.e., hidden) heat of fusion." But the kinetic theory accounts most clearly for its disappearance. If the molecules are farther apart and freer to move in the liquid than in the solid state, work is required to overcome the force of cohesion and to break them away from their molecular moorings. The heat of fusion is undoubtedly utilized in doing this work and is stored in the system as energy, for it is given out when the substance again solidifies. We have seen that temperature depends upon the average linear kinetic energy of the molecules. During fusion the temperature does not change, so we infer that the average linear kinetic energy of the molecules does not change. Therefore the heat of fusion must be stored as **potential energy.**

279. Freezing is the passing of a substance from the liquid to the solid state. After freezing begins, the temperature remains constant until the entire mass is solid. This temperature is called the **freezing point.** For pure, crystalline substances the freezing point is the same as the melting point; and usually the former is more easily determined with accuracy.

The curve plotted while a substance cools and freezes, using times as abscissas and temperatures of the substance as ordinates, is called the *cooling curve* (Fig. 238). Cooling curves are much

used in metallurgy and in the calibration of temperature measuring devices.

If carefully cooled without jarring, many substances may be cooled below the freezing point without solidifying. Water, for

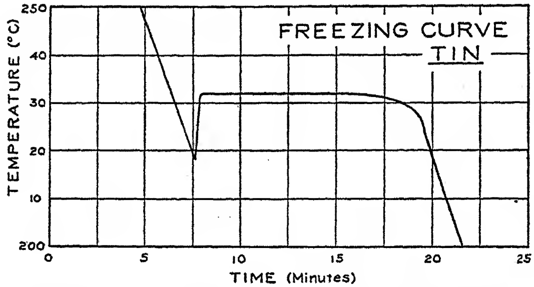


FIG. 238. Cooling Curve of Tin. (Courtesy of Professor Vance White)

example, may be cooled as much as 10°C below the freezing point without freezing. But if jarred, or if a small crystal of ice is dropped in, it freezes at once with the evolution of its heat of fusion.

280. Change of volume during freezing. Most substances contract on freezing, as the kinetic theory would lead us to expect, but water, cast iron, and bismuth are exceptions.

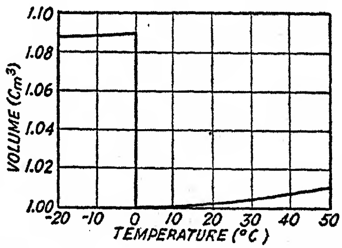


FIG. 239. Change of Volume during Freezing

Water expands about 9% on freezing, and on further cooling the ice contracts somewhat, the maximum volume being at 0°C (see Fig. 239). This accounts for the fact that water pipes often break while thawing that did not break on freezing. The ice may form plugs in the pipe, and if melting takes place between two of these plugs, the pipe is most likely to break just as the ice melts, since the volume is a maximum at that time.

An interesting fact in this connection is that, other circumstances being the same, water that has been heated freezes more quickly than water that has not been heated. The former has lost its occluded air, which gives a cushioning effect. Hence hot water

pipes often freeze and burst when cold water pipes beside them do not.

Cast iron makes sharp castings because, on freezing (1230°C), it expands and fills the mold snugly.

281. Effect of pressure on change of state. Pressure aids the state in which the volume is least.

Hence, if a substance expands on melting, increase of pressure will tend to keep it from melting, and the temperature must be raised higher than the normal melting point to cause melting; i.e., increase of pressure raises its melting point. Similarly, if a substance contracts on melting, increase of pressure lowers the melting point.

282. Regelation. Ice affords a good example of the lowering of the melting (freezing) point. Let a loop of thin wire be placed around a block of ice about 3 in. square in cross section, and a mass of 2 or 3 kg suspended from the loop, as shown in Fig. 240.

The wire being very thin, the weight causes considerable increase of pressure beneath the wire. This lowers the melting point—since ice contracts on melting—below the temperature of the ice, and consequently the ice melts under the wire. The water thus produced, not being able to support the load, flows around to the upper side of the wire, where, the pressure being relieved, it again freezes and gives out its heat of fusion, most of which passes through the wire (a good conductor) and aids the melting below the wire.

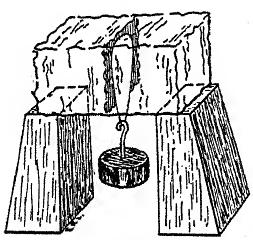


FIG. 240. Regelation

Thus the wire gradually cuts its way through the ice; but freezing takes place behind the wire, so that when the loop finally cuts out, the block is as solid as at first.

This phenomenon of refreezing when the pressure is reduced was discovered by Faraday and is called **regelation.** Often it may be observed in snow that is barely frozen. When one steps in such snow, the increase of pressure under the shoe lowers the freezing point below the temperature of the snow, so that it melts. When the foot is raised, the excess pressure being removed, the water

freezes again—but this time as ice. Hence the footprints remain molded in ice.

The change of melting point with pressure is usually small; e.g., the melting point of ice (freezing point of water) is lowered only 0.0072°C by an increase in pressure of one atmosphere, so that for most purposes the change may be neglected.

283. Vaporization. We know that most liquids, if left in an open dish, gradually disappear. This is in accord with the kinetic theory, for the molecules having the highest speeds might be expected to break away from the liquid surface occasionally and pass into the space above. There, having complete freedom of movement and relatively few collisions compared to the number in the liquid state, these molecules are in the state of a gas.

Vaporization is the passing of a substance into the state of a vapor, or gas. It may occur in three ways:

1. **Sublimation** is the changing of a solid directly into a gas without passing through the liquid state. (See Sec. 289.)
2. **Evaporation** is the vaporization of a liquid from its surface only. It takes place quietly at all temperatures.
3. **Boiling,** or ebullition, is the vaporization of a liquid when bubbles of vapor, forming throughout the body of the liquid, rise and break at the surface.

284. Boiling point. If we continue to add heat after a substance has melted, its temperature will rise till boiling begins. Its temperature will then remain constant, if the substance is pure, until all the liquid has passed into the vapor state.

The boiling point at a given pressure is the temperature at which boiling takes place. Since the bubbles break at the surface, the vapor pressure inside the bubbles must be at least equal to the total pressure above the liquid, whether that is due to the vapor alone or to a mixture of the vapor and other gases.

For pure substances, the temperature of the vapor just above the liquid is the same as that of the boiling liquid. Impurities in general tend to raise the boiling point, but the temperature of the vapor just above the liquid is still that of the pure solvent.

Since all substances expand on vaporizing, the general principle of Sec. 281 predicts the fact that **an increase of pressure raises the boiling point, and a decrease of pressure lowers it.**

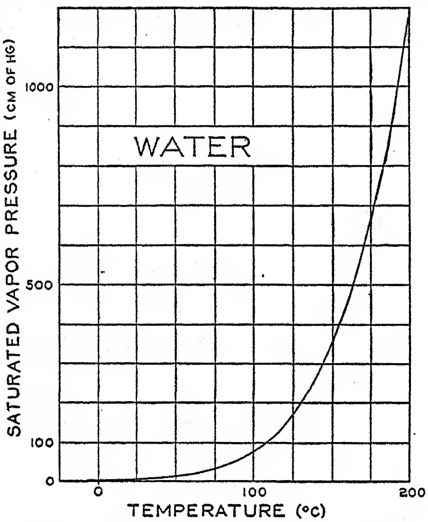


FIG. 241. Vapor Pressure Curve of Water

TABLE OF BOILING POINTS *
(At 1 standard atmosphere)

Helium	−267 °C
Hydrogen	−252.6
Nitrogen	−195
Oxygen	−182.7
Ammonia	−33.6
Water	100
Mercury	357.3
Sulphur	444.7
Iron	2450
Platinum	3910

The variation of the boiling point of water with pressure is shown in Fig. 241. From this it is seen that **the boiling point of water is 100°C (= 212°F) only when the pressure is 1 standard atmosphere.**

The lowering of the boiling point of water by reducing the pressure is shown in a striking way by the following experiment. With the stopcock of Fig. 242 open, the water in the flask *F* is first brought to boiling. The thermometer will then read 100°C if the barometer reads 76 cm of mercury. Let the burner then be removed from beneath the flask and the stopcock closed. Boiling will at once cease, and the water may be allowed to cool for 5 or 10 minutes until the temperature is 60° or 70°C. If the connection to the vacuum pump is then opened and the pump started, the water may be made to boil vigorously again **by merely reducing the pressure.**

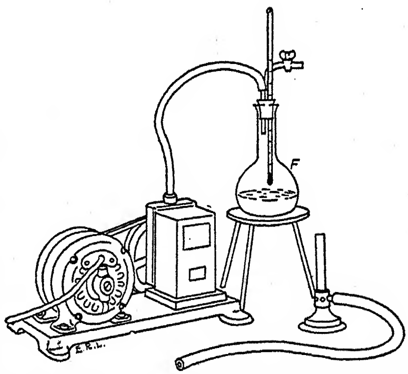


FIG. 242. Boiling at Reduced Pressure

On Pike's Peak the barometer usually stands at about 46 cm, and the boiling point is therefore about 188°F. With this low boiling point, many cooking processes are slowed up enor-

* From *Hand Book of Physics and Chemistry.*

mously, if not made impossible. The difficulty is overcome by the use of pressure cookers (Fig. 243). The relief valve on these cookers may be adjusted so that the contents will boil at any desired temperature.

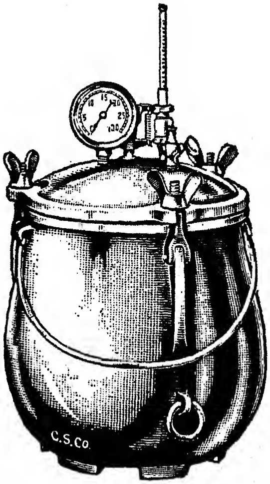


FIG. 243. Pressure Cooker. (Courtesy Central Scientific Co.)

285. Freezing and boiling points of solutions. In the case of a solution, the freezing point is lower and the boiling point is higher than for a pure solvent, on account of the attraction of the molecules of the solute for those of the solvent.

Raoult's law states that **the depression of the freezing point and the elevation of the boiling point of a solution are proportional to the molecular concentration of the solute,** except in the case of acids, bases, and salts, or of a reaction between the solute and the solvent.

Thus, 1 gram-molecular weight of methyl alcohol (32 gm) in 1 liter of water lowers its freezing point to −1.86°C; and the same depression is produced by 1 gram-molecular mass of any other solute with the exceptions mentioned.

Similarly, 1 gram-molecular mass of any non-volatile solute (say, 342 gm of cane sugar) in 1 liter of water raises the boiling point by 0.52°C; and the same elevation is produced by 1 gram-molecular weight of any solute with the exceptions named.

286. Heat of vaporization. Although we continue to add heat at a constant rate to a substance while it is vaporizing, its temperature does not increase. The heat must therefore go to increase the **potential** energy of the substance; for if its kinetic energy changed, its temperature also would change, as is indicated in Sec. 269. That the heat is stored in the substance is evidenced by the fact that it is given out again when the substance condenses, i.e., returns to the liquid state. We conclude, therefore, that this "latent heat," as it was formerly called, is utilized in overcoming the mutual attractions of the molecules and pushing them apart the relatively great distances which separate them in the gas state.

The amount of heat required to change a unit mass of a sub-

stance from the liquid to the gas state (or phase) without change of temperature is called the heat of vaporization of that substance.

This varies with the temperature at which vaporization takes place. For *water*, Henning gives the following relation: *

$$L = 598.8 - .599t \tag{236}$$

where L is the heat of vaporization in 15° calories, and t is the temperature of boiling in degrees C.

At a boiling point of 100°C, this formula gives 538.9 cal/gm as the heat of vaporization of water. A more recent value by Henning (1909) is 538.7 cal/gm.

Heats of vaporization and of fusion are readily obtained experimentally by an obvious application of the method of mixtures (Sec. 274).

287. Saturated and superheated vapor. If a liquid is in a closed vessel, the molecules having the highest speeds escape into the space above the liquid where they dart about in all directions, impinging on the sides of the containing vessel and occasionally colliding with one another. As the liquid surface is one boundary of the gas space, many of these escaped molecules, in their random motion, will dive back into the liquid.

As evaporation proceeds, a condition is finally reached when as many molecules per second dive back into the liquid as evaporate from it. Thereafter the amounts of liquid and vapor remain unchanged. The two phases (or states) of the substance are said to be in **equilibrium**; and the vapor is said to be saturated.

A saturated vapor is a vapor which is in equilibrium with its liquid.

A vapor produced in a closed vessel in contact with its liquid will always become saturated after a time, if there is sufficient liquid, and it will then exert a pressure which will depend only upon the temperature. That is, **at a given temperature a saturated vapor produces a definite pressure** which is characteristic of the substance. This pressure is called its saturated vapor pressure, or vapor tension; and when plotted against the corresponding temperature it gives the saturation curve (Fig. 241).† The pres-

* For fuller discussion, see J. H. Poynting and J. J. Thomson, *Text-book of Physics*, Vol. III: *Heat* (Philadelphia, J. B. Lippincott Company, 1913), p. 181.

† This is also the boiling-point curve, since, in order for the bubbles to break at

sures are those observed when there is no other substance present in the vessel.

If there had been some other gas such as air above the liquid (chemical action being excluded, of course), the other gas or gases would only slow up the process of evaporation. The vapor would still become gradually saturated and exert the same vapor pressure as if no other gas were present. Then the total pressure would finally be the original pressure of the other gas plus the saturated vapor pressure of the liquid. This is in accord with Dalton's law of partial pressures and is easily demonstrated as follows.

At room temperature, say 20°C, let a thin bulb containing water be suspended in a flask of air at atmospheric pressure b, the two mercury columns of the manometer being therefore at the same height; and let the flask be tightly closed (Fig. 244). If the bulb is now broken by being shaken off the hook, the water will gradually vaporize and the pressure will increase until the outer mercury column stands 1.75 cm higher than the inner one, that being the saturated vapor pressure of water at 20°C. Since the manometer is open to the atmosphere, the total pressure P within the flask is then:

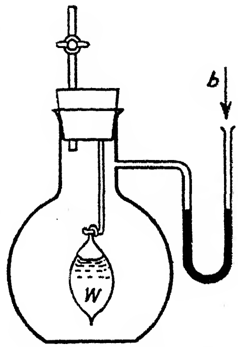


FIG. 244. Determination of Saturated Vapor Pressure

$$P = b + 1.75 \text{ cm of Hg.}$$

If all the liquid is evaporated in a closed vessel so that nothing but vapor remains, or if some of the vapor is removed to another vessel **out of contact** with its liquid, and is further heated, the vapor is then said to be superheated.

A vapor is superheated when its temperature is higher than that of saturated vapor at the same pressure.

Superheated vapors obey the gas laws much more exactly than do saturated vapors. If we have in a vessel a saturated vapor without any liquid, and if we increase the pressure, keeping the temperature constant, some of the vapor will condense and the rest will remain saturated. On the other hand, if we decrease

the surface, the saturated vapor pressure within the bubbles must be at least equal to the total pressure above the liquid, whether that pressure is due to the vapor alone or to a mixture of the vapor and other gases.

the pressure (as by increasing the volume), the vapor will be superheated.

288. The triple point. If a mixture of ice and water is maintained at 0°C and 1 std. atm. pressure, the amounts of solid and liquid present will remain unchanged. The solid and liquid phases (states) are said to be in equilibrium. We have seen in the preceding section that a liquid may be in equilibrium with its saturated vapor, as may also a solid.

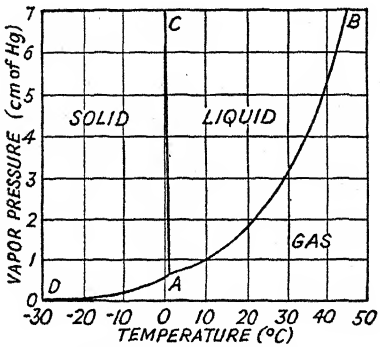


FIG. 245. Triple Point

In Fig. 245 are plotted three curves: AB gives the temperatures and corresponding pressures for which water is in equilibrium with its vapor (boiling-point curve); AC gives the temperatures and corresponding pressures for which ice is in equilibrium with water (melting-point curve); and AD gives the temperatures and corresponding pressures for which ice is in equilibrium with water vapor (sublimation curve).

These three curves meet in a point A, the triple point. The **triple point** is the point on a temperature-pressure diagram showing the temperature and pressure at which all three phases of a substance are in equilibrium at the same time.

289. Sublimation. Sublimation is the passing of a solid directly into the gas state without going through the liquid state. From Fig. 245 it is seen that this will occur at any temperature below that of the triple point when the vapor pressure of the substance in the surrounding space is less than its saturated vapor pressure for that temperature, as is shown by the sublimation curve DA.

For example, the triple point of water corresponds to a pressure of 0.46 cm of Hg and a temperature of 0.0076°C. At −2°C, say, the curve shows that the saturated vapor pressure of ice is 0.39 cm of Hg. Hence, if the pressure of water vapor in the space surrounding the ice is not as great as 0.39 cm of Hg, the ice will continue to sublime until its vapor pressure reaches that value. Then equilibrium will be established and sublimation will stop. This

accounts for "clothes freezing dry" when the temperature is below 0.0076°C and the atmosphere is very dry. In the reverse process, hoar frost forms directly from the water vapor in the atmosphere without passing through the liquid state.

Iodine, camphor, and naphthalene (moth balls) sublime at ordinary room temperatures. Solid carbon dioxide gets its trade name, "dry ice," from the fact that it always sublimes under ordinary conditions and hence never wets neighboring bodies. The triple point for carbon dioxide corresponds to a pressure of 5.11 atm. and a temperature of −56.4°C. As generally furnished, dry ice has a temperature of about −76°C, which is far below its triple point. The pressure of its saturated vapor at that temperature is 1.19 atm., whereas the pressure of carbon dioxide vapor normally in the atmosphere is negligible. Hence solid carbon dioxide always sublimes unless constrained in a gastight vessel capable of sustaining considerable pressure.

Heat of sublimation is the amount of heat required to change the unit mass of a substance from the solid to the gas state without change of temperature.

290. Humidity. Water vapor is one of the normal constituents of atmospheric air at altitudes of less than 10 mi, but the amount in a given volume of air varies greatly at different times and places.

The process of measuring the humidity, or the amount of moisture in the air, is called **hygrometry** from the Greek word, *hygros*, meaning "moist."

Absolute humidity is defined as the number of grams of water vapor per cubic meter actually in the air at a given time. It is measured by means of a chemical hygrometer (Fig. 246). A measured volume of air is passed through drying tubes containing $CaCl_2$ or P_2O_5 until thoroughly dry, i.e., until the tubes cease to gain in mass.

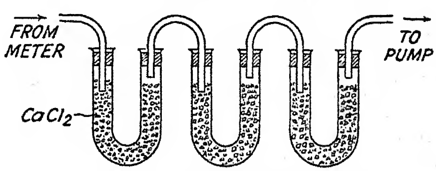


FIG. 246. Absolute Hygrometer

The increase in the mass of the drying system is the mass of water that has been abstracted from the air; and this mass, divided by the volume of air used, gives the absolute humidity.

The amount of water vapor that a given volume can hold depends upon the temperature: the higher the temperature, the more it will hold, and vice versa. **When any volume of air contains all the vapor it can hold at that temperature, it is said to be saturated.** Strictly speaking, it is the space that is saturated.

If moisture-laden air is sufficiently cooled, it will finally become saturated; then vapor will begin to condense as dew.

The dew point is the temperature at which the moisture in the atmosphere begins to condense.

If a mass of moist air is suddenly chilled below the dew point, the excess moisture is condensed and precipitated as rain. When rain drops are cooled below the freezing point, hail is formed. If the temperature is below the triple point, snow is formed in the upper atmosphere or hoar frost on bodies near the earth's surface. That is, snow and hoar frost freeze directly from the vapor state without going through the liquid state.

Generally speaking, we are not so much concerned with absolute as with relative humidity.

Relative humidity is defined as the ratio of the mass of moisture actually in a given volume of air at any time to the mass that the given volume could hold if saturated at that temperature. Or, relative humidity is the ratio of the absolute humidity at a given time to the absolute humidity if the air were saturated at that temperature. It is therefore a fraction and is usually expressed in percentage.

Relative humidity plays a most important role in daily life, for it is relative rather than absolute humidity that makes us uncomfortable on the damp warm days that we call "sultry." The habitat of plants and animals is contingent largely upon the average relative humidity.

Until recently it was thought that the most healthful humidity was about 50%. However, the following excerpt from a report of a committee of the American Medical Association, appointed to study air conditioning, seems to disprove this belief: "Under ordinary indoor conditions during the heating season, variations of humidity are relatively unimportant as far as warmth and comfort are concerned, and from the standpoint of health there are no data to prove artificial humidification is necessary. . . . No one disputes the injurious effect of low humidities to furni-

ture, but the argument about health has little foundation in fact." *

Relative humidity is a determining factor in the location of many industries. Leaf tobacco can be handled only in a very moist atmosphere. The fact that the average relative humidity of the New England states was suitable for the spinning and weaving of cotton was one of the chief reasons that the industry first developed in those states. Today, however, humidity may be maintained at any desired value by automatic humidifiers.

Relative humidity is commonly measured by one of four methods:

1. By determining the absolute humidity with a **chemical hygrometer** under the given conditions, and again for saturated air at the same temperature, then dividing the former by the latter in accord with the definition.

Fig. 247. Wet and Dry Bulb Hygrometer. (Courtesy Central Scientific Co.)

2. By means of the **wet-and-dry-bulb hygrometer.** This consists of two accurate thermometers mounted side by side on a vertical board, as in Fig. 247. The bulb of one of these is bare, while that of the other is covered with a porous cotton wick, the lower end of which dips into a reservoir of water. Evaporation from the wick cools the bulb inside the wick. The rate of evaporation depends upon the relative humidity of the air. If the relative humidity is high, the air is nearly saturated; hence it takes up water slowly and the cooling effect is slight. When the air is very dry, evaporation is rapid and cooling great. The bare bulb, however, is not affected by the humidity. Hence the difference of the temperatures given by the dry and the wet bulb thermometers is indicative of the relative humidity of the atmosphere. For best results, a current of air should be blown across the bulbs at a speed of 6 or 8 ft/sec.

The instrument does not read directly. Observations are made of the temperature on the dry bulb, the difference of temperature of the two bulbs, and of the barometer. With these values,

* C. F. Yaglou, *Journal of the American Medical Association*, CXVIII (May 15, 1937), 1708.

the relative humidity is found from tables prepared by method 1 or from curves (Fig. 247a) plotted from such tables.

3. By means of the **dew-point hygrometer** (Fig. 248), which is the device most generally used. Air is pumped through ether contained in the polished nickel vessel N. The evaporation of the

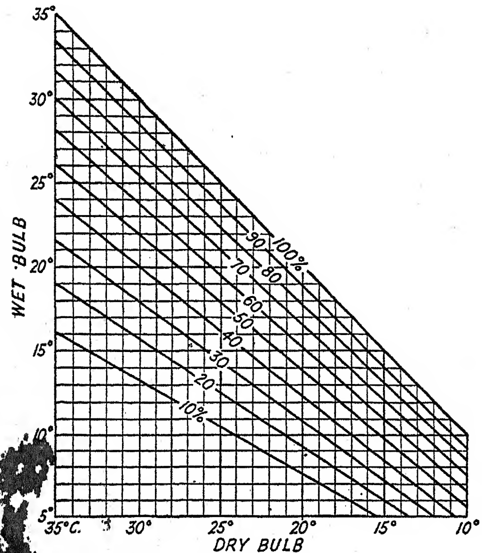


FIG. 247a. Curve for Wet and Dry Bulb Hygrometer. (Courtesy of Professor J. M. Cork)

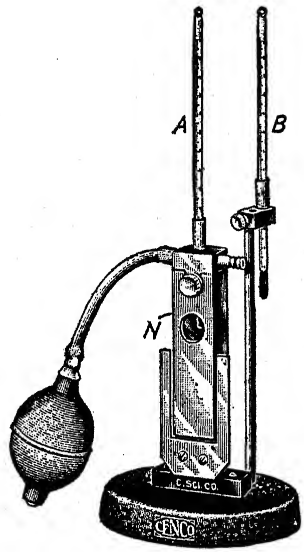


FIG. 248. Dew Point Hygrometer. (Courtesy Central Scientific Co.)

ether quickly cools N and its contents, so that moisture begins to condense on the outside. The average of the temperature at which the moisture appears and the temperature at which it disappears when the pumping is stopped is taken as the dew point.

If we assume that the mass of water vapor in a given volume of air at any time is proportional to the pressure which that vapor exerts, it follows that, approximately,

$$\text{Relative humidity} = \frac{p_d}{p_t} \tag{237}$$

where p_d is the pressure of saturated water vapor at the dew point, and

p_t is the pressure of saturated water vapor at the temperature of the atmosphere as given by thermometer B.

4. By means of the **hair hygrometer.** This utilizes the well-known property that hair has of elongating in a moist atmosphere. The device is arranged so that a number of strands of human hair *H* (Fig. 249) are kept taut by a spring. When the humidity increases, the elongation of the hair is taken up by the spring and causes a hand to move over a graduated dial or drum. The instrument must be calibrated by one of the other methods.

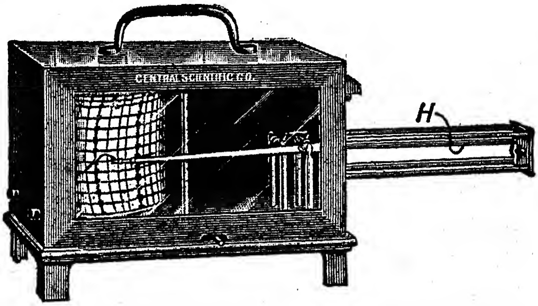


Fig. 249. Hair Hygrometer. (Courtesy Central Scientific Co.)

291. Critical state. Before 1863, oxygen, hydrogen, and nitrogen were called permanent gases because no method was known by which they could be liquefied. In that year, Dr. Thomas Andrews carried out on carbon dioxide a series of experiments which led to a clear understanding of why former attempts to liquefy these gases had failed.

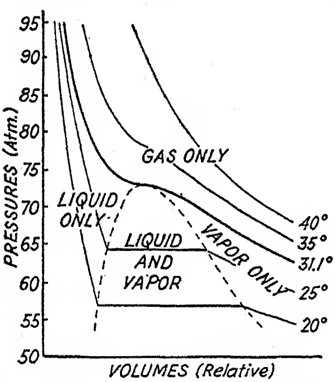


Fig. 250. Critical State

Figure 250 shows the results of Andrews' experiments. These curves, plotted with volumes as abscissas and pressures as ordinates, are called **isothermals** because for any one curve the temperature was held constant.

If 1 gm of the gas at, say, 25°C, is compressed, the pressure increases as the volume decreases until the gas begins to liquefy. The pressure then remains constant (curve horizontal) until all the gas is liquid. After that, the liquid being only slightly compressible, the pressure rises very rapidly.

For any temperature up to a certain value the isothermal curve will show this horizontal portion representing the condition when part of the substance is gas and part liquid. But as the experi-

ment is made at successively higher temperatures, a value is finally found (31.1°C for CO_2) at which the horizontal portion of the curve reduces to a single point. At the conditions represented by this point liquid and vapor have the same density. The entire mass changes from all-gas to all-liquid as the values of pressure and volume represented by this point are passed through.

The **critical point** is the point whose coordinates are the volume and pressure at which a vapor and its liquid have the same density.

The specific volume for the critical point is called the **critical volume.**

The pressure for the critical point is called the **critical pressure.**

The temperature of the isothermal curve on which the critical point occurs is called the **critical temperature.**

The critical temperature is defined as the highest temperature at which a gas can be liquefied by pressure alone. That is, above the critical temperature a substance exists only as a gas, and it must be cooled below the critical temperature before it can be liquefied.

On this basis we distinguish between a gas and a vapor: above its critical temperature a gaseous substance is called a *gas;* below its critical temperature, a *vapor.*

From the following table it is clear that the so-called "permanent" gases could not be liquefied until methods of producing very low temperatures had been developed.

Critical Temperatures

Air	−140
Ammonia	+130
Helium	−268°C
Hydrogen	−234
Nitrogen	−146
Oxygen	−118
Water	+365

292. Refrigeration. The production of low temperatures is called *refrigeration.*

One of the earliest methods of refrigeration was by means of **freezing mixtures.** Of these, the most familiar is the mixture of crushed ice, or snow, and common salt (NaCl). In order to pass into the liquid state, the ice requires heat, as does the salt on dissolving. This heat is taken from the mixture itself at the expense

of its temperature. In the proportion of 1 part of salt to 3 of ice, this mixture will fall from −1° to −21.3°C, which Fahrenheit is said to have considered the lowest temperature attainable and hence to have chosen as the zero of his scale of temperature.

Similarly, a mixture of 3 parts of calcium chloride ($CaCl_2 6H_2O$) crystals to 2 parts of crushed ice yields a temperature of −55°C, if both are originally at 0°C. Carbon dioxide snow (dry ice) and ether produce a temperature of −80°C.

Freezing mixtures, however, are relatively expensive and troublesome. Where long-continued production of cooling is required, the principle invariably employed is that of *evaporation.*

Evaporation is a cooling process. This fact is attested by the familiar experience that we feel cool when water or alcohol evaporates from the skin. It is readily accounted for by the kinetic theory of matter. Since the molecules that escape as vapor are the ones with the highest speeds, those that are left in the liquid have the lower speeds. Hence the mean kinetic energy of the molecules of the liquid is lowered, and therefore its temperature. Or, to put it another way: in order to vaporize, the molecules must acquire their heat of vaporization from somewhere and take it from the substance itself at the expense of its own temperature.

An excellent example of this method is the production of carbon dioxide snow. A tank of liquid CO_2 is turned with its outlet downward and clamped in a stand, as shown in Fig. 251. A bag of loosely woven material is tied over the nozzle, and the liquid is allowed to escape into it. Here part of the liquid evaporates, taking its heat of vaporization (33 cal/gm) from the remaining liquid so rapidly that this liquid is frozen to snow-like crystals, "carbon dioxide snow." In commerce, this snow is compressed into cakes at a pressure of 600 to 800 lb/in.^{2} and sold as **"dry ice."** It has several advantages: its temperature is −76°C; its heat of sublimation is large (153 cal/gm); it sublimes instead of melting, and consequently it does not "wet" the container.

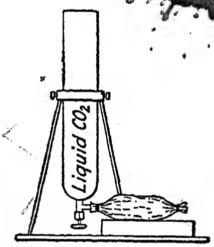


FIG. 251. Making Solid Carbon Dioxide

293. Refrigerating machines. The ordinary ice plant for the manufacture of ice from water employs the same principle, that

evaporation is a cooling process. Figure 252 shows a schematic drawing of a compression ammonia system.

Liquid ammonia at room temperature and high pressure is contained in the tank S, its quantity being indicated by the glass

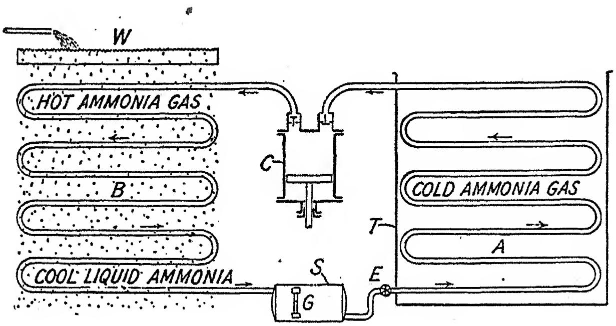


Fig. 252. Manufacture of Ice

gauge G. When the expansion valve E is opened, the liquid passes into the coils A, where it takes up its heat of vaporization (297 cal/gm) from the coils and their surrounding liquid. This liquid may be brine, which surrounds the cans containing the water to be frozen and which is pumped around through the rooms that are to be cooled. Or, as in the plate system of ice-making, the coils A are submerged in the water to be frozen and the ice is formed in sheets directly on these coils.

The cool ammonia gas at low pressure is drawn into the compressor C, from which it is expelled at high pressure (and hot) into the coils of the condenser B. Water trickles over these pipes from the serrated edges of the water trough W, and, as it evaporates, takes up its heat of vaporization from the pipes until the ammonia gas inside is cooled below its critical temperature (130°C). It then liquefies, giving out its heat of vaporization to the pipes and cooling water, and returns to the tank S. The cycle is repeated indefinitely. "Electric" refrigerators are of this type.

The **absorption type refrigerator** is shown in Fig. 252a. The refrigerant, a mixture of ammonia 30% and water 70% by weight, is heated in the generator G, whence it is elevated as in a coffee percolator to the separator S. The ammonia gas being but slightly soluble in hot water escapes in S and passes to the condenser C; while the water, cooled in the absorber A, flows back

to the generator. Condenser and absorber are provided with cooling disks.

In C the ammonia gas condenses to liquid ammonia which passes to the evaporator E. Here each gram of ammonia takes up its heat of vaporization as it evaporates, thus cooling its surroundings. It then passes back to the absorber where it is taken up by the relatively cool water in which it is highly soluble.

The total pressure in the system is maintained at several hundred pounds per square inch by the addition of hydrogen, which is practically insoluble in water, and the system is then hermetically sealed. At this pressure, condensation takes place in C at ordinary room temperatures. At the same time the vapor pressure of ammonia in E is low enough to permit evaporation readily.

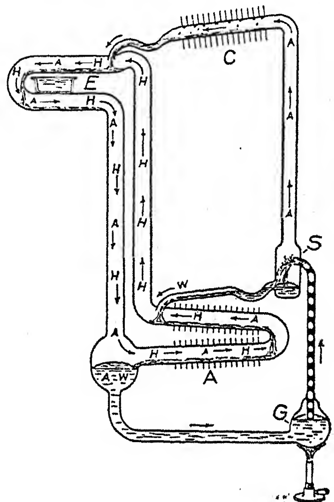


FIG. 252a. The Gas Refrigerator. (Courtesy of Servel, Inc.)

294. Liquefaction of gases. Success in liquefying all the so-called permanent gases in quantity has been achieved by what is known as the *regenerative process* due to K. P. G. von Linde (1895). This employs the principle that a gas cools on expanding, because its molecular kinetic energy is transformed into potential energy in pushing back the surrounding air and in expanding the gas itself, as in the porous plug experiment (Sec. 267).

A schematic drawing of a liquid-air machine is shown in Fig. 253. Atmospheric air, purified and dried, is drawn into the first stage C_1 of the compressor, compressed to about 250 lbs/in.^{2}, and delivered to the second stage C_2, where it is compressed to about 3000 lbs/in.^{2}, with consequent heating.

This hot air at high pressure is cooled by passing through the refrigerating bath in R, and is led to the interchanger I. The interchanger consists of three long copper pipes, number 1 being inside number 2, and number 2 inside number 3. The triple-walled tube thus made is wound into a helix (but is shown straight

in the figure for clearness) and is housed in a felt-covered cylinder H.

The air enters the interchanger by the innermost tube 1 and, on passing through the valve V_1, expands down to 250 lbs/in.^{2}. Cooled by this expansion, some of it returns between tubes 1 and 2 to compressor C_2, and in so doing it lowers the temperature of the gas that is approaching in tube 1. The rest of it expands from 250 lbs/in.^{2} down to atmospheric pressure through valve V_2. The second expansion cools this part still further; and as it passes between tubes 2 and 3, it lowers yet more the temperature of the gas in 2, finally passing out to the atmosphere.

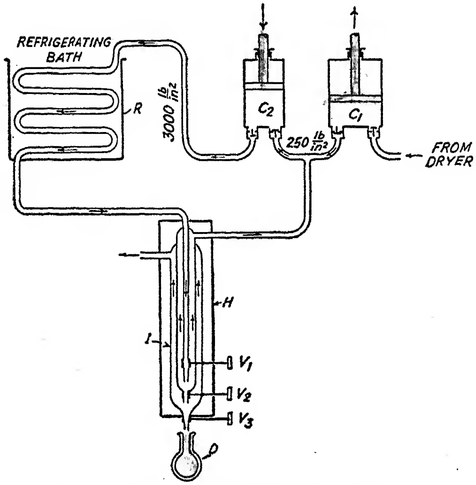


FIG. 253. Liquefaction of Air

This process of successive, or regenerative, cooling is continued until the air is brought below its critical temperature (−140°C), after which it begins to liquefy. The boiling point of liquid air is −191°C at a pressure of one atmosphere. It must therefore be further cooled to this temperature before it will remain liquid for any length of time. When this has been done, valve V_3 is opened and the clear liquid drips into the Dewar flask D (see Sec. 304). Eight hundred cubic feet of ordinary air produce about 1 ft^3 of liquid air.

At first the liquefied air contains both liquid nitrogen and liquid oxygen. But at standard atmospheric pressure the boiling point of nitrogen is −195.5°C, while that of oxygen is −182.7°C. Hence the nitrogen evaporates first, leaving, after a few hours, practically pure oxygen.

By suitable adaptations of the regenerative method, hydrogen (boiling point −252.8°C) was liquefied by Sir James Dewar in 1898, and was solidified by him the following year. Using liquid hydrogen to cool compressed helium, which was then allowed to

expand, the Dutch physicist H. K. Onnes succeeded, in 1908, in reducing this, the last of the "permanent gases," to liquid form. The boiling point of helium at atmospheric pressure is −268.9°C, and by allowing liquid helium to evaporate in a vacuum, Onnes produced a temperature which he estimated, from the vapor pressure, to be only 0.82°C above absolute zero.

In 1926, Keesom produced solid helium at a temperature of −271.9°C and a pressure of 26 atmospheres.

PROBLEMS

1. Find the heat required to change 400 gm of ice at −10°C to water at 60°C.

2. Taking the heat of fusion of ice to be 80 cal/gm, compute its value in Btu per lb.

3. What is the average rate at which heat penetrates into a refrigerator if 20 lb of ice in it are melted in 24 hours?

4. If 200 gm of ice at −8°C are placed in 80 gm of water at 25°C, find the temperature of the mixture and the amount of water when equilibrium is established.

5. How much steam at 100° will be required to melt 1 kg of ice at −10°C?

6. How much heat will be required to change 10 lb of ice at 20°F into steam at 1 atmosphere pressure?

7. How much ice at 0°C must be mixed with 1 gal of boiling water at 100°C to make a mixture at 50°C?

8. Steam at 212°F is turned into a cast-iron radiator whose mass is 150 lb and whose temperature is 32°F until its temperature becomes 75°F. How much steam is condensed?

9. How much steam at 100°C must be added to 40 gm of ice at −20°C, so that the final temperature of the mixture will be 70°C?

10. If the heat of vaporization of ammonia is 297 cal/gm, how much liquid ammonia at 0°C must be evaporated to make 1 kg of ice from water at 16°C?

11. How much steam at 100°C must be added to 40 gm of ice at −8°C so that the final temperature of the mixture will be 50°C?

12. A copper ball is heated in a steam bath at a pressure of 76 cm of mercury and is then dropped into a hollow block of ice. If 8 gm of ice melt, what is the mass of the ball?

13. What will be the final temperature when 150 gm of ice at 0°C, 400 gm of water at 20°C, and 40 gm of steam at 100°C are mixed?

14. If 500 gm of superheated steam at 180°C are mixed with 2500 gm of ice at −2°C, what is the final temperature of the mixture? (Specific heat of the steam = 0.48 cal/gm; of ice = 0.50 cal/gm.)

CHAPTER XXII

TRANSFER OF HEAT

295. Modes of transfer. Heat, which we have seen to be molecular energy, is transferred from place to place in three ways:

1. By conduction.
2. By convection.
3. By radiation.

Radiation is perhaps the most important, since it is in this manner that heat comes to us from the sun.

296. Conduction is the transfer of heat from molecule to molecule of the same body or of different bodies.

If we assume that the temperature of a body depends on the mean linear kinetic energy of its molecules, then molecules at higher temperatures would transfer some energy to those of lower temperature when they collide; and hence heat would flow along a body from the region of higher to the region of lower temperature. This is just what we know takes place, and it would seem that the rate of flow of heat by conduction would be different for different kinds of molecules, or materials. The experiment of Fig. 254 verifies this.

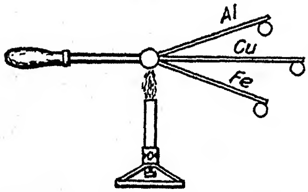


Fig. 254. Relative Conductivities

Rods of aluminum, copper, and iron, having the same dimensions, are fastened in a central sphere of copper which is provided with a handle. If balls of beeswax of equal masses are stuck to the lower sides of the rods at the outer ends, and the central sphere is heated over a Bunsen flame, heat will be transferred to the wax by conduction along the rods.

The balls will melt loose first from the copper, then from the aluminum, and last from the iron. Hence we conclude that copper is a better conductor than aluminum, and aluminum better than iron.

Silver is the best conductor known, although copper is almost as good. Dry air is one of the poorest conductors of heat.

297. Thermal conductivity. Just as we should expect, the amount of heat H transmitted by conduction varies directly as the cross-sectional area A of the body, the difference of temperature $(t_2 - t_1)$ between two cross sections, and the time of flow T; and inversely as the distance d between the two cross sections. (See Fig. 255.)

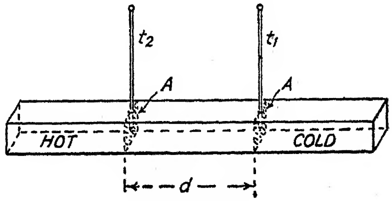


FIG. 255

Algebraically,

$$H \propto \frac{A(t_2 - t_1)}{d} T$$

or

$$H = \frac{kA(t_2 - t_1)}{d} T. \tag{238}$$

The coefficient k depends upon the choice of units and is called the *thermal conductivity*. To enable us to define k in words, we make all the factors of the right side of the equation unity except k, thus:

$$\mathrm{H_{cal}} = \frac{k\ 1\ \mathrm{cm^2}\ 1°\mathrm{C}}{1\ \mathrm{cm}}\ 1\ \mathrm{sec.}$$

From this it is seen that: **The thermal conductivity k of a substance is the number of calories transmitted in 1 second between opposite faces of a cubic centimeter of the substance when the temperature difference between those faces is 1°C.** A similar definition would apply if all the units were taken in the British system.

For most pure metals the thermal conductivity decreases when the temperature increases; but the reverse is true of alloys and of most electrical insulators, e.g., glass. Liquids and gases in general are very poor conductors of heat.

On comparing these values with the electrical conductivities of the same substances, it will be found that, in general, good heat conductors are also good conductors of electricity, and vice versa.

TABLE OF THERMAL CONDUCTIVITIES
(In cal cm^{-1} sec^{-1}°C^{-1})

Substance	Temp.	k
Silver	18°C	1.006
Copper	18	0.918
Aluminum	18	0.504
Iron	18	0.161
Glass	....	0.0025
Quartz (fused)	20	0.0024
Brickwork	....	0.0015
Water	17	0.0013
Ice	....	0.004
Snow	....	0.0005
Air	0	0.00006

For pure metals, $\dfrac{k}{\gamma T}$ is approximately constant

where k is the thermal conductivity;
γ is the electrical conductivity; and
T is the absolute temperature.

298. Heat insulation. In the construction of refrigerators, cold storage rooms, and even in residences, it is important to reduce the loss of heat by conduction through the walls as much as possible. Various so-called insulating materials such as ground cork, rock wool, hair felt, etc., all of which have low thermal conductivities, are commonly used in such cases.

But still, dry air is the most effective heat insulator (except a vacuum); and the efficacy of these many materials lies in their ability to prevent convection currents and hold the air stationary in tiny compartments. For this reason, pipe covering material is made cellular by concentric layers of corrugated magnesia or asbestos paper.

The insulating value of wool for clothing is due largely to the fact that wool fibers have a slightly rough surface to which air adheres with exceptional tenacity.

299. Convection is the transfer of heat by the motion of the hot body as a whole, as when heat is carried from a furnace to upper rooms by steam or hot air. Convection plays a more prominent part in daily affairs than does conduction.

The Gulf Stream and the Trade Winds are natural convection currents, and hurricanes and tornadoes originate as convection currents over highly heated portions of the earth.

In industry, and in the ordinary household utilities, convection is employed for many useful purposes. Figure 256 illustrates the production of chimney draft. The air over the candle expands on being heated, so that its density is less than that of the cooler air outside. Hence it is buoyed upward like a piece of wood released under water, in accordance with the principle of Archimedes, the colder air descending to take its place. A circulation, or convection current, is therefore set up: hot air upward at the left, cool air downward at the right.

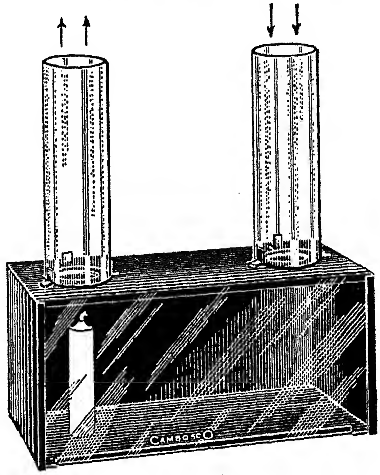


FIG. 256. Convection in Air. (Courtesy Cambridge Botanical Supply Co.)

Another illustration of convection is furnished by the hot-water tank of Fig. 257. The water, heated in the water front which forms one side of the firebox, expands with consequent decrease of density. It is therefore buoyed up by the more dense cool water, and rises to the top of the tank. If not drawn off in the service mains, it gradually cools somewhat, mixes by diffusion with the cooler water, and descends on the side opposite to the stove, thus taking part in the circulation as shown by the arrows.

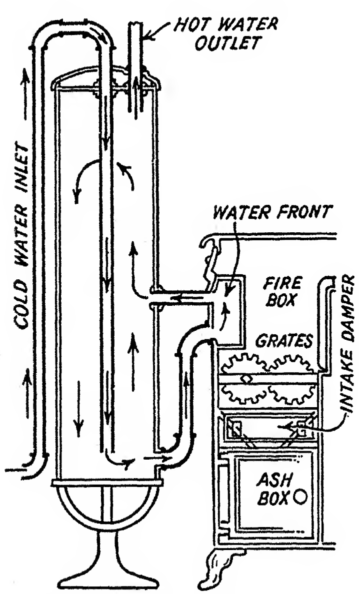


FIG. 257. Convection in Hot-Water Tank. (From Snyder's *Physics*, courtesy of Allyn and Bacon)

300. Radiation. Most of our heat comes to us from the sun, but

the space that separates us from that luminary (93,000,000 mi) is practically devoid of matter. It is therefore impossible that this heat comes to us by conduction or convection, for either of these methods would require intervening matter. Hence we say that this heat, or solar energy, is transmitted to us by waves similar to those of light. The process is called radiation, which name is also often applied to the waves themselves.

Radiation is the transmission of energy by electromagnetic waves. These waves were supposed to occur in the ether—a hypothetical, all-pervading medium proposed by James Clerk Maxwell for the accommodation of light waves. But ether waves consist of varying electric and magnetic fields, and are not material waves like those in water. According to more recent theory the idea of the ether is unnecessary, but we shall retain it for the present as a convenient concept.

The ether waves themselves are not heat; they merely transmit energy. When they fall upon matter, the waves set the molecules of the matter into vibration, and the energy has then become heat. This is attested by the fact that the empty space between the earth and the sun, which is practically devoid of matter, is not heated, although these waves pass through it. The energy of the waves in transit is often called "radiant energy," "radiation," "radiance," or "radiant heat." But it is not heat in the ordinary sense until it falls upon matter and sets the molecules in motion.

301. Radiant heat and light. The radiation, or radiant heat, that reaches us from a hot stove, like that from the sun, has all the properties of light (it may be reflected, refracted, polarized, etc.) except that it is not visible; i.e., its wave lengths are too long to produce in the eye the sensation of sight.

Bodies that transmit heat waves are said to be **diathermanous;** those that do not, **athermanous.** It is surprising to see that hard rubber, though opaque to light, transmits long heat waves well, whereas clear glass is quite opaque to them.

Therein lies the secret of the **greenhouse effect.** A large part of the radiation of the sun that reaches the surface of the earth consists of the visible waves and the short infrared, because the very short (ultraviolet) waves and the very long (infrared, or heat) waves are largely absorbed by our atmosphere. The glass roof of

a greenhouse, being very transparent to these visible waves, admits them, and their energy is converted into heat when they strike the objects inside. As these objects warm up, they themselves become radiators; but their temperatures not being high, they emit only very long waves. Glass is athermanous to heat waves longer than about 1.5μ, hence these longer waves radiated by the bodies in the greenhouse cannot get out. A greenhouse therefore acts as a **heat trap,** and the temperature inside will become higher than the outside temperature in direct sunlight. For the same reason, an ordinary thermometer reads too high when its bulb is in direct sunlight.

The carbon dioxide and water vapor in the lower layers of the atmosphere serve in the same way to trap the long heat waves radiated by bodies on the earth. Consequently the temperature near the earth's surface is much higher than that in the upper regions of the atmosphere.

302. Ideal black body. An ideal black body is defined as a body that absorbs all the radiation that falls upon it. In the following section it will be seen that an ideal black body, when sufficiently heated, emits radiation of all wave lengths. The type of radiation from an ideal black body depends only upon its absolute temperature.

No perfect black body is known, but a surface coated with lamp black or bismuth black serves as a very good one. Lummer and Pringsheim have found * that the blackened interior of a uniformly heated enclosure (Fig. 257a) having only one opening, and that very small, is **practically a perfect black body** for experimental purposes; for the fraction of the radiation entering the hole that will come back out, after repeated reflections, is negligible. The interiors of most large furnaces and of long heated tubes closed at one end closely approximate black body conditions if kept at a steady temperature.

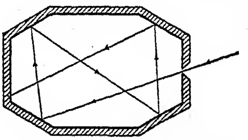


FIG. 257a. Black Body of Lummer and Pringsheim.

303. Radiating, or emissive, power. The *radiating*, or *emissive, power* of a surface at a given temperature is the amount of energy

* J. M. Cork, *Heat* (New York, John Wiley & Sons, 1933), p. 156.

radiated by it per square centimeter per second when at that temperature.

Although radiation originates in the volume elements of a body, radiating power is greatly affected by the nature of the surface, as may be shown by means of Leslie's cube (Fig. 258). The faces of the cube c are finished in four ways: one is highly polished metal; one is painted white; another red; and another black. If this cube is filled with boiling water and placed at a definite distance from a thermopile T connected to a galvanometer G, it will be found that the deflections are different when the different faces are presented to the thermopile. The highly polished metal surface has the least radiating power, and the black surface the greatest. A rough surface radiates better than one that is polished.

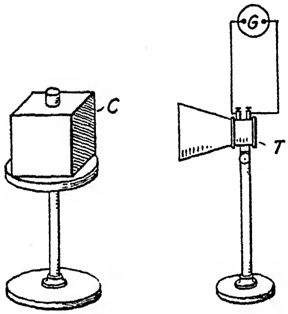


FIG. 258. Leslie's Cube

The **emissivity** of a surface is the ratio of the emissive power of the surface to the emissive power of an ideal black surface under the same conditions.

Absorptivity,* formerly called absorbing power, is the ratio of the energy absorbed by a surface at a given temperature to the total energy incident upon it. From the definition of an ideal black body, its absorptivity is unity.

If we fill the cube of Fig. 258 with liquid air instead of boiling water, the galvanometer will then deflect in the opposite direction, for the former hot junction will now be the cold junction. The surface of the cube facing the thermopile is then absorbing heat radiations from the thermopile; and it will be found that the black surface has the greatest absorptivity and the highly polished one, the least.

While the emissivities and absorptivities of different surfaces differ greatly, the emissivity of any one surface equals its absorptivity. If this were not true, we could polish one set of junctions of a thermopile and blacken the other set, place the pile in a room, and have perpetual motion. For the black surface, being a

* Not to be confused with "absorption coefficient." See F. K. Richtmyer, *Introduction to Modern Physics* (New York, McGraw-Hill Book Company, 1928), p. 186.

better absorber than the polished surface, would become warmer; the thermopile would develop an emf which would operate a sufficiently delicate motor; and enough of these devices would give us unlimited power. This is obviously absurd on its face; as a matter of fact, both sets of junctions stay at room temperature, neither rising above the other. Clearly, then, the black surface must radiate as rapidly as it absorbs; and the polished surface must radiate as rapidly as it absorbs. Hence, **for any one surface the emissivity at a given temperature equals the absorptivity at that temperature.**

Reflectivity is the ratio of the energy reflected by a surface at a given temperature to the total radiation incident upon the surface. If a body is so thick that no radiation passes through it, the energy that is not absorbed must be reflected.

From the foregoing facts we are able to state the **law of Kirchhoff and Stewart: Bodies which are good reflectors are poor radiators and vice versa;** and **the wave lengths a body absorbs are the same that it radiates.** For this reason, teakettles and hot water bottles are made of polished metal in order to reduce the loss of heat by radiation.

304. The Dewar flask. This most useful device for keeping substances hot or cold was invented by Sir James Dewar in 1892, and illustrates well the application of the preceding principles.

The flask is a double-walled bottle (Fig. 259), the space between the walls being exhausted to a very high vacuum. This prevents loss of heat by conduction or convection (except up the neck) since there is no matter between the walls to conduct or to convey heat.

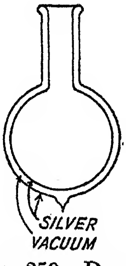


FIG. 259. Dewar Flask

The outer surface of the inner wall and the inner surface of the outer wall are both made mirrors by a deposit of silver. These excellent reflectors reduce to a minimum the loss of heat by radiation.

Dewar flasks are usually made of glass, but they are now available also in polished steel. Comparatively cheap forms of this flask are found on the market under various trade names.

Liquid air is generally kept in the best grade of Dewar flask, the top being closed by a loose plug of cotton wool. In summer a liter

(quart) of liquid air will keep for two or three days in a flask of this kind.

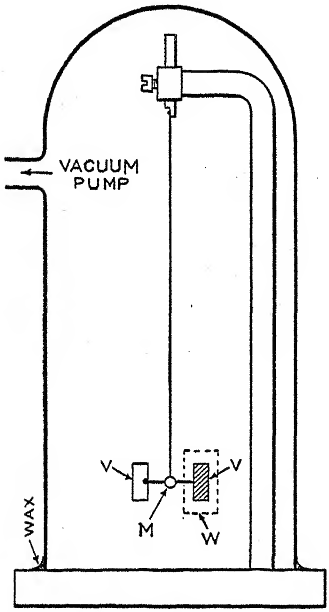


FIG. 260. Nichols' Radiometer

305. Radiometers. Radiation is usually measured by means of a radiometer consisting of a vacuum thermocouple on the hot junction of which is fastened a small, thin "receiver" of silver or copper whose surface opposite to the junction is coated with bismuth or platinum black. The glass bulb containing the thermocouple must be provided with a window of *rock salt* or *sylvine* or *fluorite*, which are very diathermanous.

The deflections of a sensitive galvanometer connected to such a vacuum thermocouple are proportional to the radiant energy per second falling upon the receiver. W. W. Coblentz of the Bureau of Standards has constructed radiometers of this type capable of registering the radiation from a candle at a distance of 53 mi (Sec. 448).

Another type of radiometer (devised by Edward L. Nichols) consists of two tiny vanes of mica or glass mounted on a light horizontal bar which is suspended at its middle point by a delicate fiber of quartz and carries a very small mirror. One side of each vane is blackened, the other side not (Fig. 260).

This moving system is suspended in a partly exhausted tube having a diathermanous window. The radiation falls upon the blackened side of one vane, the other vane being completely shielded from the radiation; and the deflection is proportional to the intensity of the incident radiation.

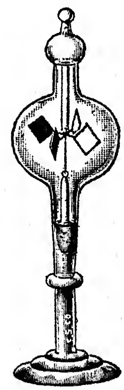


FIG. 261. Crooke's Radiometer. (Courtesy Central Scientific Co.)

This type of radiometer is constructed also with four or more vanes so as to make the radiation motor (Fig. 261) often seen spin-

ning merrily in the sunlight in display windows. The action is as follows.

Radiation is absorbed more rapidly by the blackened surfaces than by those not blackened. Hence the temperature of the blackened sides will rise slightly higher than that of the other sides. Molecules of the rarefied air in the bulb are heated more when they come in contact with the blackened surfaces than when they strike the unblackened ones. Hence they rebound with greater speeds from the black surfaces, and give those surfaces an impulse in the opposite direction. The wheel therefore turns with the unblackened surface leading.

306. Prévost's theory of exchanges. Recognizing that the rate at which energy is radiated by a body **depends only upon its temperature** and **the nature of its surface,** Pierre Prévost formulated the following theory: All bodies above the temperature of absolute zero are radiating heat all the time, and all bodies are receiving heat all the time; and when the temperature of a body is constant it is receiving heat at the same rate as that at which it is radiating.

307. The Stefan-Boltzmann law. In 1879, Stefan obtained from experimental data the law that **the total heat *H* of all wave lengths radiated per square centimeter per second by a black body is proportional to the fourth power of its absolute temperature *T*.** Subsequently Boltzmann derived the same law from electromagnetic theory. Stated algebraically,

$$H = CT^4. \tag{239}$$

The Stefan-Boltzmann constant C has the value: 5.72×10^{-5} erg cm^{-2} deg^{-4} sec^{-1} when cgs units are used.

308. Rate of radiation. If a certain body is at the temperature $T°K$ and its surrounding body or bodies have the temperature $T_1°K$, then the body is radiating to its surroundings an amount of heat H per cm^2 per sec and is absorbing from its surroundings an amount H_1 per cm^2 per sec. If both the body and its surrounding body or bodies are **black bodies,** then by the Stefan-Boltzmann law:

$$H = CT^4$$

and

$$H_1 = CT_1^4.$$

Hence the net rate of radiation of the body is:

$$H - H_1 = C(T^4 - T_1^4). \tag{240}$$

309. Planck's law of radiation. The Stefan-Boltzmann law gives the rate of radiation of energy of all wave lengths together from a black body. In order to find how much energy was radiated in each wave length band of width $d\lambda$, i.e., in the wave length range from λ to $\lambda + d\lambda$, Lummer and Pringsheim passed the radiation of a black body through a spectrometer, thus resolving it into its constituent wave lengths. Measuring the energy in each wave length band separately, they obtained curves similar to those of Fig. 262.

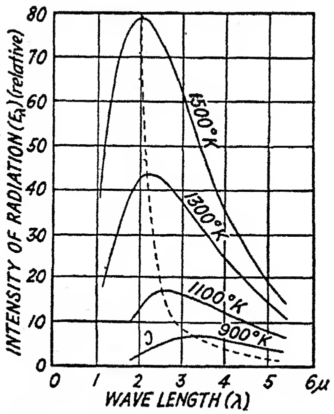


FIG. 262. Distribution of Black Body Radiation

Of various attempts to fit an equation to these curves, the most successful was that of Max Planck of the University of Berlin. His equation gives, with a high degree of accuracy, the distribution among the different wave lengths of the energy radiated by an experimental black body. It is therefore known as **Planck's spectral distribution law:**

$$J_\lambda = \frac{c_1}{\lambda^5\left(\epsilon^{\frac{c_2}{\lambda T}} - 1\right)} \tag{241}$$

where J_λ is the energy radiated per sec per cm² of surface per unit of $d\lambda$ in the interval from λ to $\lambda + d\lambda$;

- T is the absolute temperature of the black body in °K;
- λ is the mean wave length in the interval;
- ϵ is the base of the Naperian system of logarithms;
- c_1 is the first radiation constant and depends upon the choice of units; and
- c_2 is the second radiation constant upon which chiefly depends the shape of the curve.

For cgs units,*

$$c_1 = 3.697 \times 10^{-5} \text{ erg cm}^2 \text{ sec}^{-1}$$
$$c_2 = 1.438 \text{ cm deg.}$$

The product $J_\lambda d\lambda$ represents the energy radiated per cm^2 per sec in the wave length band whose width is $d\lambda$. Hence the area under the curve represents the total energy of all wave lengths radiated per sec from each cm^2 of the surface of the black body.

310. The quantum theory. In attempting to derive his spectral distribution law as a rational equation, Planck found that the classical electromagnetic theory of Maxwell led to obvious inconsistency. He therefore abandoned Maxwell's theory, which requires that energy must be radiated in a continuous, uninterrupted stream, and evolved for the purpose the entirely new quantum theory (1900).

The quantum theory assumes that energy is radiated in discrete pieces, or packages, each containing an amount of energy called a "quantum," or "photon."

$$\text{A quantum of energy} \equiv h\nu \tag{242}$$

where h is a universal constant, called Planck's constant, and ν is the frequency of the waves of the radiation.

In cgs units, $h = 6.62 \times 10^{-27}$ erg sec; and ν is equal the velocity of the radiation (= that of light) divided by the wave length λ.

$$\nu = \frac{3 \times 10^{10} \text{ cm/sec}}{\lambda \text{ cm}}. \tag{243}$$

From this it is seen that ν is large when λ is small, and vice versa. Consequently, the amount of energy represented by **one quantum is different for radiations of different wave lengths,** being greater for short wave lengths and less for long wave lengths.

The quantum theory accounts satisfactorily for the spectral distribution of radiation from a black body, for the photoelectric effect (Einstein), and for spectral series (Bohr), and is the cornerstone of much of modern physics. However, it does not account for interference, polarization, and similar phenomena easily explained by the electromagnetic wave theory.

* R. T. Birge, in the *Rev. Mod. Physics*, Vol. 13, No. 4 (Oct., 1941).

In 1924, Louis de Broglie suggested that the ultimate particles of matter might be essentially wave-like in nature, which was confirmed in 1927 when Davisson and Germer showed that electrons could be diffracted like light. From such considerations, the new "wave mechanics" of Schroedinger and Dirac and the "quantum mechanics" of Heisenberg have evolved. Their success thus far makes it hopeful that they will ultimately reconcile the discrepancies between the wave and quantum theories of radiation.

311. Wien's displacement law. The curves of Fig. 262 show that as the absolute temperature T increases, the wave length (λ_{max}) in which the maximum energy is radiated is displaced toward shorter wave lengths, as is shown by the dotted curve. This is true for all bodies.

A familiar illustration of this shift to shorter wave lengths is the fact that with a small current in the filament of an electric lamp the light emitted is quite red. As the current (and temperature) is increased, the light becomes yellow and finally quite white on account of the increasing amount of blue (short) radiation emitted at the higher temperatures.

Wilhelm Wien has shown that, for a black body,

$$\lambda_{max}T = \text{Constant.} \tag{244}$$

which is **Wien's displacement law for black body radiation.** It may be deduced from Planck's law, and is confirmed by experiment.

312. Black body temperature. An ideal black body not only absorbs all radiations that fall upon it, but at any given temperature it radiates at the maximum possible rate in each wave length. That is, at a given temperature no body can radiate in a given wave length more rapidly than a black body at the same temperature.

If any non-black body is radiating at a certain rate in a given wave length, its **black body temperature is the temperature that an ideal black body must have in order to radiate in that wave length at the same rate as the non-black body.**

Since non-black bodies are less efficient radiators than black bodies, it follows that the **true temperature** of a non-black body **is always greater** than its black body temperature. For example,

a piece of iron whose true temperature is 1200°C will register a black body temperature of only 1140°C when tested with an optical pyrometer using red light. If green light were used, the black body temperature for that wave length would be different from that for the red.*

313. The radiation pyrometer. The platinum resistance thermometer is not reliable above 800°C, nor the platinum platinum-rhodium thermocouple above 1500°C. Temperatures higher than these are measured by some form of radiation pyrometer. By such means the temperature of the surface of the sun is found to be about 6000°*K*.

Radiation thermometers are of two types:

1. Those called **radiation pyrometers,** in which the total radiation, or a definite part of it, is focused upon and absorbed by some kind of receiver which actuates an indicator.

2. Those called **optical pyrometers,** in which a photometric match is made between the image of the hot body and a filament or other device whose temperature (and color) can be controlled in a definite way.

An excellent example of the first type is the **total radiation pyrometer** of Fery (Fig. 263).

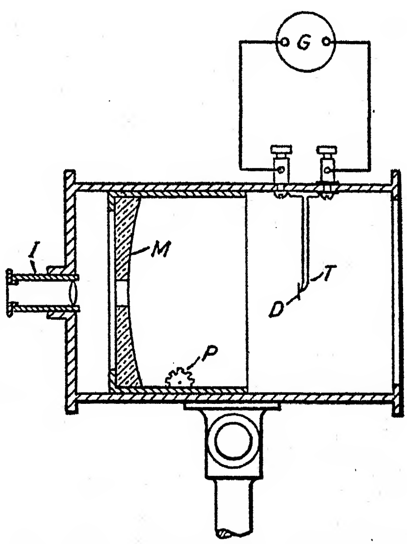


FIG. 263. Total Radiation Pyrometer

The observer, sighting through the eyepiece *I*, focuses the image of the hot body by means of the gold-plated, concave mirror *M* on the blackened surface of a small silver disk *D*. To the back of this disk is attached a thermocouple *T* which is connected to a sensitive galvanometer *G*.

If the image of the hot body completely covers the disk, the readings of the galvanometer, which indicate the temperature, are

* Burgess and LeChatelier, *Measurement of High Temperatures* (New York, John Wiley and Sons, 1912), p. 242.

practically independent of the distance of the pyrometer from the hot body.

Since the total radiation of all wave lengths is focused on the disk D, the scale of temperature for this radiation pyrometer is defined by the Stefan-Boltzmann law.

The instrument is usually calibrated by observations on a black body at known temperatures. The galvanometer scale may have these temperatures marked directly upon it, or a curve may be drawn by plotting the temperatures against the readings of G in microvolts. In either case the instrument gives the black body temperature of the observed body when corrections are made for the energy reflected by the body.

314. The optical pyrometer. This also is a radiation pyrometer, but employs only a narrow band, or range, of wave lengths and

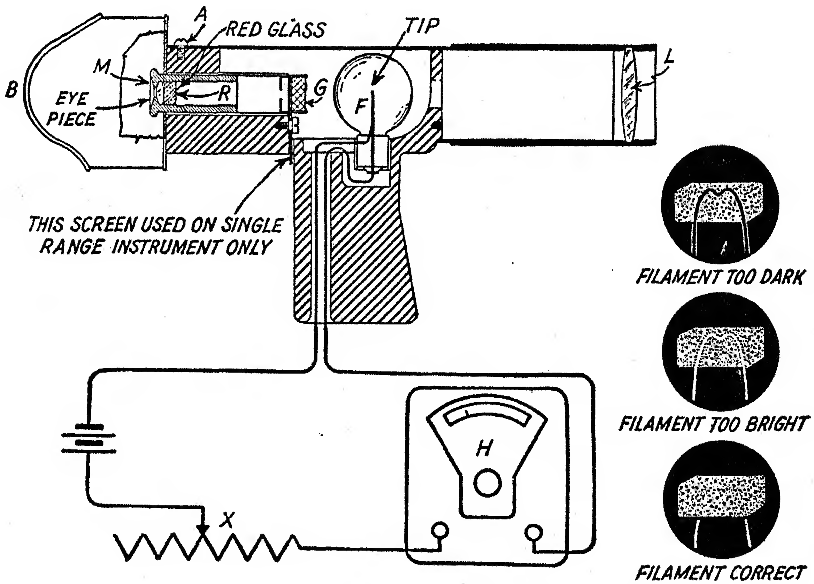


FIG. 264. Optical Pyrometer. (Courtesy of the Leeds and Northrup Co.)

not the total radiation. Hence the scale of temperature for pyrometers of this type is defined by Planck's law of radiation (or the simplified form of this known as Wien's law). The effective wave length is usually about 6500 Å.

An optical pyrometer of the **disappearing filament** type is shown in Fig. 264. The image of the hot body is brought to a focus by the lens L in the plane of the filament F. The current in F is varied by means of the rheostat X until the dip at the center of the filament disappears against the image of the hot body as a background. The reading of the milliammeter H then corresponds to the temperature of the hot body.

The instrument is calibrated against a black body at known temperatures and consequently gives the black body temperature of the body observed, if correction is made for surface reflection.

PROBLEMS*

1. How much heat per hour will pass through a window whose area is 1.5 m^2 and thickness 3 mm, when the temperature is 20°C inside and 0°C outside?

2. A piece of asbestos board 20 cm square and 1.5 cm thick transmits 120,000 cal per hr when one surface is at 100°C and the other at 350°C. Compute the coefficient of thermal conductivity for the board.

3. In British units the coefficient of thermal conductivity is expressed in Btu per hr per ft^2 per in. thickness for a temperature difference of 1°F. If the coefficient for brick in cgs units is 0.0015, compute its value in British units.

4. If the thermal conductivity of pine wood is 0.98 Btu per hr per ft^2 per in. thickness per 1°F difference of temperature, how much heat will pass through a door 3×7 ft whose average thickness is 0.8 in. in 24 hr, when the inside and outside temperatures are 70° and 10°F, respectively?

5. How much heat per hour will pass through a glass window 2×3.5 ft if the glass is 1/10 in. thick, when the temperature is −10°F outside and 70°F inside?

6. A house has 400 ft^2 of brick wall 13 in. thick. If the conductivity of brick is 4.5 in British units, how much heat is lost per hour by conduction, when the thermometer outside is at 32°F and that inside is at 70°F?

7. If, in the above house, there are 360 ft^2 of window glass 1/8 in. thick, how much heat per hour is lost through the glass?

8. Assuming perfect circulation of water, how much heat will pass per hour through a wrought-iron crown sheet 2.5 ft^2 and 3/8 in. thick when the water is at 212°F and the temperature in the firebox is 700°F?

9. In a meat market, the refrigerator showcase has 90 ft^2 of exposed glass 1/4 in. thick. If the inside temperature is 35°F and the outside temperature 75°F, how much heat passes into the case in 24 hr?

* A thin layer of air or water next to the surface of a conductor may make the actual temperature at the surface quite different from that in the neighboring medium. This accounts for the seeming great difference between theoretical and practical results sometimes observed in this type of problem.

CHAPTER XXIII

THERMODYNAMICS

315. Thermodynamics is that branch of the science of heat which treats of the relations between heat and mechanical work. It was founded by the French physicist, Sadi Carnot (1796–1832), in an endeavor to increase the efficiency of the steam engine.

316. The first law of thermodynamics is merely the extension of the law of conservation of energy to include heat, or molecular energy. It may be stated as follows: **Heat may be converted into work, or work into heat; and the amount of work is proportional to the heat.** Algebraically,

$$W \propto H$$

$$W = JH \qquad (245)$$

where W is the work in *work units;*

H is the corresponding heat in *heat units;* and

J is a constant, depending upon the choice of units, called the mechanical equivalent of heat.

The ordinary steam and gas engines are, strictly speaking, **heat engines** and convert heat into work. Hero's reaction turbine (Fig. 265) illustrates this. Water is contained in the spherical boiler, to which **heat only** is supplied. The heat changes the water into steam which, discharging through the tangential nozzles N, N, reacts against them and turns the turbine as shown.

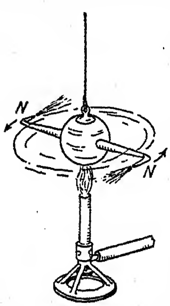


FIG. 265. Hero's Engine

Even more familiar are conversions of work into heat. Particles from a grinding wheel are heated to incandescence. When a hole is bored into hard wood or metal, both the drill and the material are heated. If a volatile liquid such as alcohol is placed in a brass tube (Fig. 266) and this is turned rapidly between the halves of a wooden brake, the heat developed will vaporize the alcohol and cause the cork to be blown out.

317. The mechanical equivalent of heat. The constant of proportionality J in Eq. (245) is called the "mechanical equivalent of heat," or "Joule's equivalent," after James Prescott Joule, who was one of the first to make a determination of its value.

In words, **the mechanical equivalent of heat is the number of units of mechanical work which are equivalent to one unit of heat.**

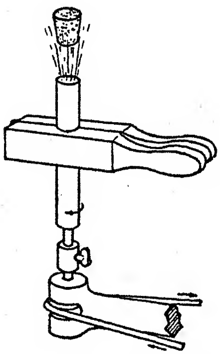


FIG. 266. A Friction Brake

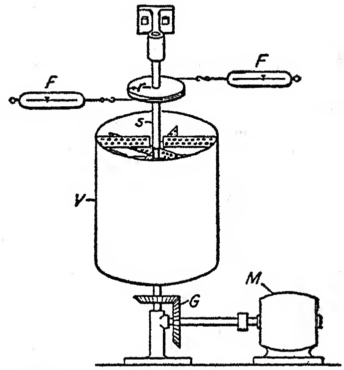


FIG. 267. Rowland's Apparatus

The classical determination of this constant was made by Henry A. Rowland in 1879 at the Johns Hopkins University, using the apparatus illustrated (in principle) in Fig. 267.

A known mass of water, placed in vessel V, is stirred by the two sets of perforated paddles, one of which is fixed to V and the other to the shaft S. V is rotated by the motor M and gears G. The shaft S and its attached paddles are held stationary relative to the earth by the spring balances F, F. The friction between the water and the paddles causes the work done in stirring to be converted into heat; in consequence, the temperature of the water rises, being measured by thermometers not shown.

Let N be the total number of revolutions;
$t_2 - t_1$ be the rise in temperature of water and calorimeter;
M be the mass of water;
c be the mean specific heat of water;
C be the thermal capacity of calorimeter, paddles, etc.; and
F be the reading of the spring balances.

The work W done by the paddles on the water is, by Sec. 77,

$$W = T\theta = (2Fr)2\pi N$$

and the heat H produced is, by Sec. 273,

$$H = (cM + C)(t_2 - t_1).$$

Hence,

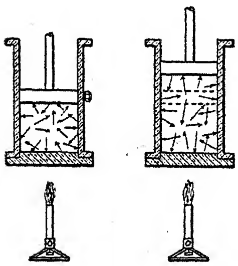
$$J = \frac{W}{H} = \frac{4\pi NFr}{(cM + C)(t_2 - t_1)}. \tag{246}$$

The accepted values of J are now: *

$$J = 4.185 \text{ joules/calorie}$$
$$= 778 \text{ ft-lb/Btu}.$$

In making this research, Rowland found that his results for J differed when the runs were made over different temperature ranges. From this he concluded that the specific heat of water was **not constant** as it had always been assumed to be; but that it was different at different temperatures. A special investigation was therefore made to determine the variation of the specific heat of water with temperature, and the results are shown in Fig. 234.

318. Two specific heats of a gas. The amount of heat required to raise the temperature of 1 gm of a gas 1°C will depend upon how the gas is constrained. The two most important cases are:



(*a*) (*b*)

FIG. 268. Two Specific Heats of Gas

(a) When the volume is constant (c_v).

In Fig. 268a, let 1 gm of gas be confined in a cylinder whose piston is fastened by a setscrew, so that the volume is constant. The heat applied has only *one* thing to do; i.e., to increase the kinetic energy of the molecules corresponding to a rise of 1° in temperature. The amount of heat required to do this is c_v.

(b) When the pressure is constant (c_p).

In Fig. 268b, the setscrew has been removed so that the piston may rise. When heat is applied the pressure tends to increase, and the piston being free, is pushed upward against the constant pressure due to the atmosphere and the weight of the piston. In this case, therefore, the heat applied has to do *three* things:

1. To increase the kinetic energy of the molecules an amount corresponding to a rise of 1° in temperature.

* International Critical Tables.

2. To do external work in lifting the piston and pushing back the atmosphere.

3. To do internal work in separating the molecules of the gas itself, except in the case of a perfect gas.

The amount of heat required to do this is c_p. It would seem obvious from these facts that

$$c_p > c_v$$

and this is confirmed by experiment.

The *ratio* of the specific heats of a gas is an important quantity. For air,

$$\gamma \equiv \frac{c_p}{c_v} = 1.40. \tag{247}$$

It may be readily shown * that for a perfect gas

$$c_p - c_v = \frac{R}{J} \tag{248}$$

where R is the gas constant for unit mass and J is the mechanical equivalent of heat.

319. The *P-V* diagram. Let a gas at constant pressure p expand behind a piston from a volume v_1 to a volume v_2 (Fig. 269).

If A is the area of the piston, the total force F acting on the piston is $F = pA$; and if the length of the stroke is L, the work done is:

$$W \equiv FL \cos \phi = pAL.$$

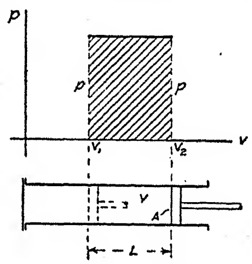


FIG. 269. Pressure-Volume Diagram I

But AL is the change of volume, $(v_2 - v_1) \equiv \Delta v$, so that

$$W = p(v_2 - v_1) = p\Delta v. \tag{249}$$

But $p(v_2 - v_1)$ is the crosshatched "area under the curve," i.e., the area bounded by the curve, the ordinates of its end points, and the v-axis.

Therefore the area under a p-v curve is proportional to the work done by the substance during the changes represented by

* Edser, *Heat for Advanced Students* (New York, The Macmillan Co., 1936), p. 295.

the curve.* This relation is true whether p is constant during the change or not. For suppose the values of p vary as shown by the curve MN of Fig. 270. The area under the curve may be divided into a number of narrow rectangles, each of width dv, so small that within this width p may be considered constant. Then each little rectangle represents an amount of work,

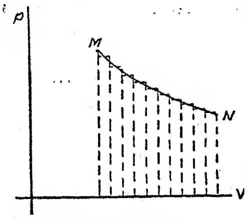


FIG. 270. Pressure-Volume Diagram II

$$dW = p\,dv$$

and the total work done during the whole change from M to N is, in the convenient symbols of the calculus,

$$W = \int_{v_1}^{v_2} p\,dv$$

which in proper units is also the total area under the curve.

320. Reversible and irreversible processes. Cycle. A **reversible process,** or change, is an **ideal** one in which equilibrium conditions prevail at every instant. That is, it must take place with infinite slowness, and there must be no friction or other cause whereby energy is disseminated so that it cannot be returned to its original source and condition without the expenditure of additional energy.

Strictly speaking, there are no reversible processes in nature, although many approach reversibility sufficiently closely for practical purposes. Thus, slow evaporation of a substance in a well-insulated vessel is practically reversible, because if the temperature of the heater is lowered ever so little, some of the vapor will condense, returning its heat to the source, and this can be continued until both substance and heater are in their original condition. Similarly, the slow compression of a fine steel spring is a reversible process.

All processes that are not reversible are said to be **irreversible.** The latter include all processes involving friction and electrical resistance. Of a different type is the following example. Consider two 1-liter flasks connected by an airtight tube and stopcock. Let one contain air at 3 atmospheres pressure and the other at 1 atmos-

* The factor of proportionality will depend upon the scales used in laying off the p's and v's on the diagram.

phere. If the stopcock is opened, air will flow quickly from the former to the latter flask; there will be temporary turbulence; but a state of thermal equilibrium will finally be attained with the pressure 2 atmospheres in each flask. Even if the flasks are well insulated so that there is no appreciable loss of energy, the gases cannot be restored to their original condition without work from an outside source. The process described is irreversible.

A **cycle** is a succession of changes, or processes, that at the end bring the body or system back to its initial state. A **reversible cycle** is a cycle all of whose processes are reversible.

321. Indicator diagrams. The device shown in Fig. 271, known as an indicator, automatically draws a *p-v* diagram whose abscissas and ordinates are proportional to the volumes and pressures, respectively, of the substance in the cylinder of the engine at each instant during a revolution or cycle. Such a curve for a steam engine is shown in Fig. 272, and Fig. 273 shows the cross section of the engine.

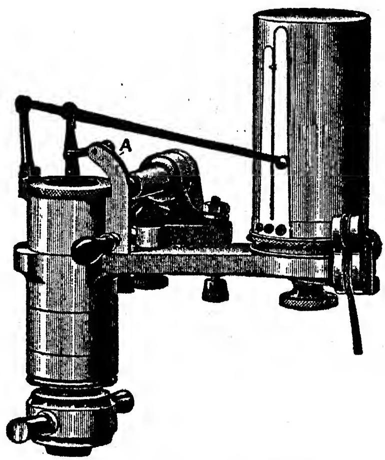


FIG. 271. Steam Engine Indicator. (Courtesy Crosby Steam Gage and Valve Co.)

Just before the piston *B* reaches the end of its travel to the left, the slide valve *d* opens the left port and permits steam from the steam chest *c* to enter behind the piston and push it to the right. While the steam is being admitted, the indicator draws the "admission line" *MN* of the indicator diagram.

At *N*, "cutoff" takes place, the valve *d* closing the port so that no more steam is admitted. The steam entrapped in the space *A* then expands, pushing the piston ahead, and the indicator draws the "expansion line" *NO*. The valve then connects the space *A* to the exhaust pipe *e*, and the pressure drops quickly almost to atmospheric. The reverse stroke then begins and the steam is pushed out of *A* by the piston, as is indicated by the line *QS*. When the piston reaches the position corresponding to *S*, the valve

again closes the left part, trapping some steam in A. This steam is compressed during the remainder of the back stroke, as is indicated by the "compression line" SM, to the original pressure at M.

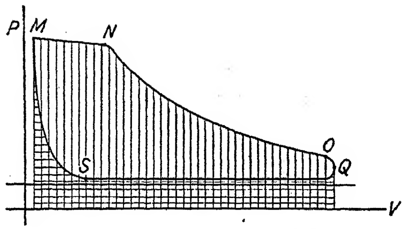


FIG. 272. Indicator Diagram

This cycle of operations is repeated over and over.

In Sec. 319, we saw that the area under a p-v curve is proportional to the work done by a substance in expanding (or being compressed) through those values.

Consequently, the area crosshatched vertically in Fig. 272 under the line $MNOQ$ represents the work done by the steam on the piston during the expansion stroke; and that crosshatched horizontally under the curve QSM represents the work done on the steam by the piston during the exhaust stroke.

Therefore, the difference of these areas, *which is that inside the diagram*, indicates the net, or effective, work done by the steam on the piston during one complete cycle.

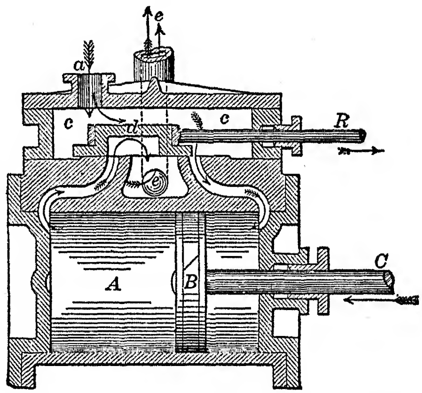


FIG. 273. Cylinder of Steam Engine. (From Carhart and Chute's *Physics*, courtesy of Allyn and Bacon)

322. Indicated horsepower. From an indicator diagram such as the above the input, or *indicated horsepower* (I.H.P.), of an engine is easily computed.

The "mean effective pressure" is obviously proportional to the average vertical width of the indicator diagram. The volume swept through per stroke is $V = AL$, where A is the area of the piston and L is the length of the stroke.

The work W_1 done in one cycle is therefore:

$$W_1 = PV = PAL$$

and if there are N cycles or power strokes per minute, the total

work done per minute is:

$$W = PALN.$$

Hence the indicated horsepower is:

$$\text{I.H.P.} = \frac{PLAN}{33{,}000} \tag{250}$$

provided that

P is the mean effective pressure in lbs/in.^{2};
L is the length of the stroke in ft;
A is the area of the piston in in.^{2}; and
N is the number of power strokes per minute.

322a. The steam turbine. The reciprocating engine has certain inherent defects: the sliding bearings are difficult to keep in adjustment, the speed is restricted to a few hundred revolutions per minute, and the entering steam is chilled by coming into contact with metal at the temperature of the exhaust steam.

The steam turbine (Fig. 273a) does not have these disadvantages. It consists of a rotating cylinder, or rotor R, having fastened radially to its surface a series of curved blades which move between a series of similar blades fixed to the housing H. The steam enters at A and is deflected by the first ring of fixed blades so that it impinges on the first ring of movable blades, driving them forward. From these it passes into the second ring of fixed blades, which deflect it against the second ring of movable blades, and so on as shown at the right side of the figure.

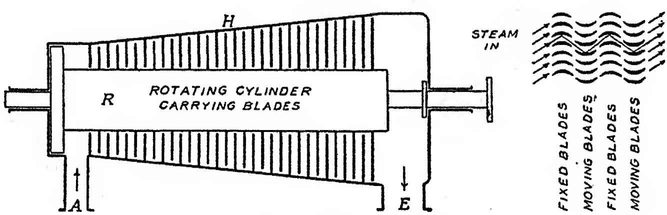


Fig. 273a. Steam Turbine

In its passage through the turbine, the steam expands and cools as its energy is imparted to the rotor; and it is exhausted through

the large pipe E. As in the case of the water turbine, the input power is PV, where P and V are the pressure and corresponding volume per second of the entering steam.

Steam turbines are made developing up to 160,000 kw on a single shaft with thermal efficiencies as high as 28% as compared to 21% for the best reciprocating engines. Still higher efficiencies are secured by using mercury and steam turbines in combination.

323. The gas engine. The gas engine is today probably the most familiar form of heat engine. The fuel, mixed with the proper amount of air, is burned inside the engine cylinder, instead of under a separate boiler as in the case of the steam engine. The heat evolved by the combustion causes the gaseous products of combustion to expand, pushing the piston outward; and thus the heat is converted into mechanical work.

Most gas engines are of the four-stroke-cycle, or "four-cycle," type. That is, a complete cycle consists of four strokes (Fig. 274), as follows:

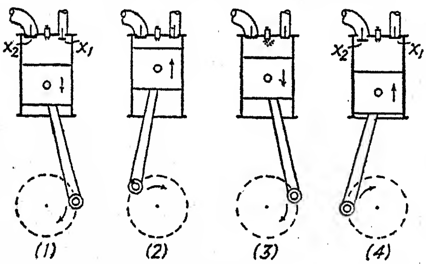


FIG. 274. Four-Stroke Cycle Gas Engine

1. Intake stroke: the explosive mixture is drawn into the cylinder through the valve x_1 as the piston descends.

2. Compression stroke: x_1 closes and the mixture is compressed as the piston rises.

3. Power stroke: the mixture is exploded by a spark just before the piston reaches the end of the compression stroke, and the piston is driven downward.

4. Exhaust stroke: the products of combustion are exhausted (driven out of the cylinder) through the valve x_2.

The cycle is then repeated, beginning again with the intake stroke.

324. The temperature-entropy diagram. The first law of thermodynamics states that work and heat are mutually convertible. We have seen that the area of a p-v diagram represents work, and it often aids analysis to have a diagram of the same type to represent heat.

When a body gains or loses heat, the change most commonly observed is that of its temperature. Temperature is accordingly taken as one of the coordinates for the desired diagram. The other coordinate that we shall have to use, in order that the area of the diagram may represent the heat gained or lost by a body, has been called by Clausius the **entropy** of the body.

Entropy is just as real a property as pressure, volume, and temperature. These, however, can be detected by the senses of touch and sight, whereas entropy cannot be detected by any of the five senses. But neither can radio waves be so detected, although we are perfectly sure of their existence. Both were first predicted by mathematical reasoning.

Experiment indicates that if the temperature of a body were reduced to absolute zero, its entropy also would become zero; but generally **we deal only with differences of entropy,** as shown in Fig. 275.

Consider a body at an absolute temperature T in a nonconducting and nonradiating vessel, and let its entropy be ϕ_1 and its heat content H_1 above those of some chosen zero condition. Plotting ϕ_1 and T on rectangular axes, this initial condition is represented by the point A in the diagram. Keeping the temperature constant, let heat be added to the body until it has the amount H_2 above that of the zero condition. Its entropy will then have increased to a new value ϕ_2 above that of the zero condition. Plotting these coordinates, the final condition of the body is represented by the point B. The quantity of heat added is then proportional to the area under the temperature-entropy curve AB; or, with properly chosen units,

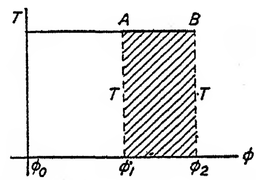


FIG. 275. Temperature-Entropy Diagram

$$H_2 - H_1 = T(\phi_2 - \phi_1)$$

$$(\phi_2 - \phi_1) \equiv \frac{H_2 - H_1}{T}. \tag{251}$$

Hence, for reversible processes, **the change of entropy of a body is defined as the ratio of the heat absorbed by the body to the absolute temperature at which the change takes place.**

If we allow the temperature to change also, the heat must be

added by infinitesimal amounts dH, so small that for a single such addition the temperature may be considered constant. The entropy also will then change by an infinitesimal amount $d\phi$, and the definition takes the most general form:

$$d\phi \equiv \frac{dH}{T}. \tag{252}$$

As an example, let us take 1 gm of water as the body, and compute its change of entropy when it passes from the solid state (ice) at 0°C to the liquid state at 100°C, the pressure being 1 std. atm throughout.

To melt the ice, its heat of fusion (79.7 cal) must be added at the constant absolute temperature (273°K). The change of entropy during melting is therefore:

$$(\phi_2 - \phi_1) \equiv \frac{H_2 - H_1}{T} = \frac{79.7}{273} = 0.292 \text{ cal}/°K.$$

After the ice has all melted, the addition of further heat causes the temperature to rise; and we shall have to find the corresponding change of entropy by adding up the little $d\phi$'s given by Eq. (252). This is the process of integration:

$$(\phi_3 - \phi_2) \equiv \int_{273}^{373} \frac{dH}{T} = \int_{273}^{373} \frac{c\,dT}{T} = 1 \times \log_e \frac{373}{273} = 0.312 \text{ cal}/°K$$

where the specific heat c of water is taken as approximately unity.

The total change of entropy of 1 gm of water at atmospheric pressure, when passing from the state of ice at 0°C to liquid water at 100°C, is the sum of these two parts, i.e.,

$$(\phi_3 - \phi_1) = 0.292 + 0.312 = 0.604 \text{ cal}/°K.$$

As will be seen later, increase of entropy is a measure of the thermodynamic degradation of energy, i.e., of its loss of availability for conversion into useful work.

325. The second law of thermodynamics. It is a matter of common observation that **heat tends of itself to pass from a hotter to a colder body.** The second law of thermodynamics postulates that it **always** tends so to do.

Thus, if we place a red-hot iron in water, experience leads us to expect the iron to cool and the water to be heated. It would not

violate the first law of thermodynamics for heat to pass from the water to the iron, causing the former to freeze and the latter to become white hot; but the second law affirms that the probability of such a transfer is zero.

The law has been stated in many ways. One of the best statements is the following, due to Clausius: **No self-sustaining, cyclical process can transfer heat from a colder to a hotter body.**

In an electric refrigerator, heat is transferred from the cold interior to the warmer exterior. But to do this, work is required by an external agent, generally a motor; and the process is not self-sustaining, because energy must be supplied to the motor or other agent from an outside source.

Heat is often referred to as the lowest form of energy, because, whereas all other forms—electrical, mechanical, chemical, etc.—are ultimately transformed 100% into heat by friction and other dissipative causes, when heat is transformed into other forms of energy, some is always wasted or rendered less available. An example is the energy rejected to the sink by the Carnot engine (Sec. 327) which is more efficient than any other engine. That energy is unavailable for use in an engine whose sink is at a temperature equal to or higher than T_1.

Analysis shows that this loss of availability of energy is measured by the increase in the entropy of the system. Such loss takes place in all actual processes in nature: they are all irreversible, and in irreversible processes, the entropy of the system increases. Hence another statement of the second law of thermodynamics is: *The entropy of the universe tends to become a maximum.*

326. Isothermal and adiabatic curves. The energy that a body has in virtue of the motion and structure of its molecules is called its **intrinsic energy.** The intrinsic energy of a body may change in various ways, but the two simplest changes are the following:

1. An **isothermal** change is one that takes place without change of temperature; i.e., T = constant.

2. An **adiabatic** change is one that takes place without the addition or withdrawal of heat *as such* from the body.

In an adiabatic change, the intrinsic energy of the body may change on account of changes of pressure, volume, and temperature. But if ΔH stands for the gain or loss of energy as heat—i.e.,

by conduction, convection, and radiation—then an adiabatic change is defined by $\Delta H = 0$. During a reversible adiabatic change, the entropy is constant; but in an irreversible adiabatic change the entropy increases.

For a perfect gas, the equations of these curves on a *p-v* diagram are as follows:

1. *Isothermal.* In the general gas law ($PV = RMT$), R and M are constant if we deal with a given mass, say, 1 lb. If the change is isothermal, T is also constant. Hence the equation of an isothermal for a perfect gas is:

$$PV = C', \text{ a constant,}$$

2. *Adiabatic.* In more advanced texts * it is shown that for a perfect gas, the equation of an adiabatic is:

$$PV^{\gamma} = C'', \text{ a constant,}$$

where γ is the ratio of c_p to c_v (Sec. 318).

Since in an adiabatic expansion no heat is added to keep up the temperature of the medium, while in an isothermal expansion heat is added, it follows that the pressure will fall more for a given change of volume when the expansion is adiabatic than when it is isothermal. Hence, **at a given point on a (*p-v*) diagram an adiabatic curve is steeper than an isothermal** for the same body. Thus, in Fig. 276b, BC is steeper than AB.

Although it is practically impossible to produce perfectly isothermal or adiabatic changes in the laboratory, these concepts are of the utmost importance in thermodynamic theory.

327. Carnot's ideal engine. Just as there is an advantage for theoretical purposes in considering an ideal transformer, a perfect gas, an ideal black body, etc., there is also an advantage in considering an ideal heat engine. Sadi Carnot, in an endeavor to improve actual heat engines, proposed such an engine in 1824, though it could not possibly be constructed.

This ideal engine (Fig. 276a) is frictionless. The piston and cylinder walls are perfect nonconductors of heat, but the cylinder head is a perfect heat conductor. R is a source, or reservoir, of heat at the absolute temperature T_2; N is a perfect nonconduct-

* Edser, *op. cit.*, p. 312.

ing shelf; and S is a sink, or receiver, of heat at the absolute temperature T_1. R and S have perfectly conducting tops and great capacity, so that their temperatures are not altered during the operation of the engine.

The heat medium used in the engine may be *anything*, but for the sake of definiteness we will call it a gas. Suppose the piston

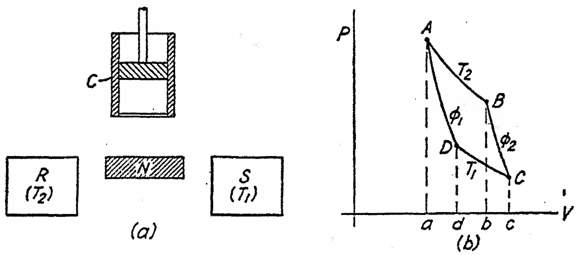


FIG. 276. Carnot's Ideal Engine and Cycle

near the bottom and the gas to have pressure, volume, and temperature p_2, v_2, T_2, respectively, represented by point A on the diagram (Fig. 276b).

The **Carnot cycle,** all the processes of which must be reversible, consists of four parts:

1. The cylinder is placed on the source R and the piston is released. As the gas expands an infinitesimal amount, its temperature decreases an infinitesimal amount. But, as the cylinder head is a perfect conductor, heat immediately flows into the gas from R, so that the expansion takes place at the constant temperature T_2, as indicated by the isothermal AB.

2. The cylinder is placed on the nonconductor N, and the gas is allowed to expand without gain or loss of heat as such, until the temperature falls to T_1, as shown by the *adiabatic* curve BC.

3. The cylinder is placed on the sink S; a hypothetical flywheel (not shown) pushes the piston downward; the gas is compressed at the constant temperature T_1, as shown by the *isothermal* curve CD, and loses heat to the sink.

4. The cylinder is then replaced on N, and the mechanism allowed to compress the gas adiabatically along DA back to its initial condition (p_2, v_2, T_2) at A; and the cycle is complete. The point D must be correctly chosen so that the adiabatic curve through D will also pass through A.

Since the slope of an adiabatic curve is greater than that of an

isothermal through the same point, the p-v diagram will enclose an area which, in accordance with Sec. 319, is proportional to the work done by the engine in executing one Carnot cycle, as this cycle is called.

If H_2 is the amount of heat received by the gas from the source, and H_1 the amount discharged into the sink, the heat converted into useful work is $H_2 - H_1$ heat units. Changing this into work units by the first law of thermodynamics:

$$\begin{aligned} &\text{Useful work got out of engine} \equiv W = J(H_2 - H_1) \\ &\text{Total work applied to engine} \qquad\quad = JH_2 \end{aligned}$$

Hence

$$\text{Efficiency of Carnot engine} = \frac{J(H_2 - H_1)}{JH_2} = \frac{H_2 - H_1}{H_2}. \tag{253}$$

This engine is representative of **reversible engines,** that is, engines that execute reversible cycles. It may be made to traverse the cycle in the reverse order $ADCBA$ by means of an external motor. In so doing an amount of heat H_1 will be withdrawn from the sink at the temperature T_1, and an amount $H_2 = H_1 + \dfrac{W}{J}$ will be delivered to the source at the higher temperature T_2, where W is the amount of work done on the engine by the motor and equals the external work done by the engine when operating on the direct cycle.

By means of the second law of thermodynamics, it may be shown * that all ideal reversible engines (i.e., having different working substances as heat media) operating between the same two temperatures have the same efficiency; and that no irreversible engine can have a higher efficiency than this when working between these temperatures. This is known as *Carnot's theorem.*

328. Kelvin's absolute thermodynamic scale of temperature. In 1848, Lord Kelvin recognized in Carnot's cycle a means of defining a scale of temperature that would be independent of any medium.

Consider two adiabatic curves A and B on a p-v diagram (Fig. 277). Let a certain mass of any medium be employed in a

* Edser, *op. cit.*, p. 336.

Carnot's engine, taking in an amount of heat H_{100} from a source at 100°C, and discharging an amount H_0 into a sink at 0°C.

The total crosshatched area would then represent the useful work done and the heat usefully employed during this cycle.

Let this same area now be divided by isothermal lines into 100 *equal* areas. We may assign to these isothermals the numbers 99, 98, 97, . . . 0, and think of the total amount of work as having been obtained from 100 little Carnot's engines each executing one of the little cycles, receiving heat from the preceding engine as a source, and discharging heat into the following engine as a sink. The isothermals that we have drawn would define a scale of temperature.

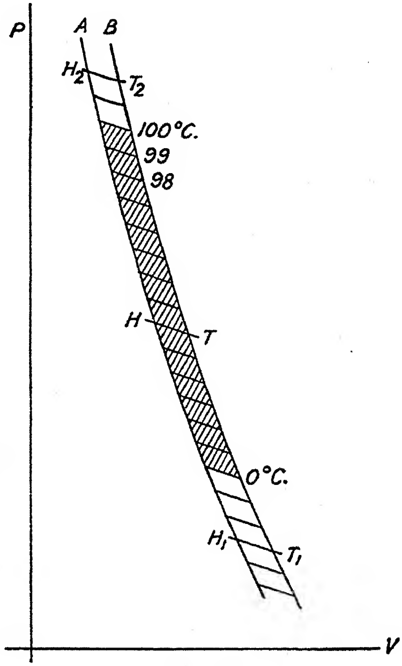


FIG. 277

Let us continue dividing the area between the two adiabatics and below 0°C by isothermal lines into little areas each equal to one of those already cut off above, until the medium has no more heat. Then no more Carnot cycles can be executed because there is no heat to be received by the next engine. With the last isothermal, we will have arrived at a natural **absolute zero** (−273.16°C) of temperature. Using this absolute zero, the scale of temperature defined by the isothermals we have drawn is **Kelvin's absolute thermodynamic scale.**

The Kelvin scale is identical with the temperature scale defined by a perfect gas,* and closely approximates the constant volume hydrogen scale. Since it is obviously impossible to construct a thermometer to indicate temperatures on the Kelvin scale, they are determined by calculation and are indicated by the symbol (°K).

That the Kelvin scale of temperature is independent of the

* *Ibid.*, p. 341.

properties of any thermometric substance is easily shown. Suppose that the amount of work represented by the area of each little cycle diagram in Fig. 277 is w work units. Then, by Sec. 316,

$$w = Jh,$$

and

$$h = \frac{w}{J}$$

is the amount of heat converted into work during each little cycle.

The number of cycles counted from absolute zero is always the same as the number T of the upper isothermal of the last cycle counted. Hence, for a Carnot engine operating between any two temperatures T_2, T_1, on the Kelvin scale,

$$H_2 = hT_2 \quad \text{and} \quad H_1 = hT_1$$

and therefore the

$$\text{Efficiency of Carnot's engine} = \frac{H_2 - H_1}{H_2} = \frac{h(T_2 - T_1)}{hT_2} = \frac{T_2 - T_1}{T_2}. \qquad (254)$$

Since the efficiency of the Carnot engine is independent of the heat medium used, it follows that **the Kelvin absolute scale of temperature is independent of any medium.**

PROBLEMS

1. Find the heat developed by the brakes of a 2500-lb car when bringing the car to rest from a speed of 45 mph on a level road.

2. Find the heat developed by the brakes of a 3000-lb car when brought to rest from a speed of 60 mph on a level road.

3. If the calorific value of coal is 14,000 Btu per lb, how much water must be stored at a height of 100 ft to have the same potential energy as that stored in 1 lb of coal?

4. A lead bullet is fired from a gun with a speed of 400 m/sec against an iron plate. If the mass of the bullet is 10 gm and 50% of the heat developed goes into the bullet, how much is its temperature raised?

5. A double-acting steam engine having a diameter of 12 in. and a stroke of 18 in. makes 400 rpm. If the mean effective pressure shown by the indicator card is 38 lb/in.², what is the input power?

6. A double-acting steam engine has a piston diameter of 10 in. and a stroke of 16 in. If the speed is 500 rpm and the mean effective pressure as shown by the indicator card is 40 lb/in.², what is the input power?

7. A steam plant uses 400 lb of coal per hr and has an over-all efficiency of 7%. What is the horsepower of the plant? (Calorific value of coal = 14,500 Btu/lb.)

8. If a steam engine has an over-all efficiency of 11% and develops 5 hp, how much heat must be supplied to it per minute?

9. A steam plant uses 500 lb of coal per hr and has an over-all efficiency of 8%. What is the horsepower of the plant? (Calorific value of coal = 14,000 Btu/lb.)

10. If the calorific value of gasoline is 20,000 Btu/lb and its specific gravity 0.72, how many gallons will be required to operate a 45-hp gasoline engine for 1 hr if its efficiency is 25%?

11. A 65-hp automobile requires 3 gal of gasoline per hr on an average run. What is its efficiency? (Calorific value of gasoline is 20,000 Btu/lb.)

12. If an automobile running at 50 mph develops 36 hp and has an engine efficiency of 18%, how many miles per gallon can it travel using gasoline having 120,000 Btu per gallon?

13. A cubic inch of water makes approximately 1 ft^3 of steam at a pressure of 1 atmosphere. How much mechanical work is done in vaporizing 1 lb of water in the open air?

14. A certain steam engine receives steam at 150°C and exhausts it to the atmosphere at 100°C. What is its maximum possible efficiency? What would its maximum efficiency be if it exhausted into a condenser at 50°C?

15. If superheated steam is used at 330°C, what would the thermodynamic efficiency be in the second part of problem 14?

16. A combination mercury-steam turbine unit takes saturated mercury vapor from a boiler at 480°C and exhausts it into a steam boiler at 240°C. The steam turbine takes steam at this temperature and discharges it into a condenser at 40°C. What is the thermodynamic efficiency of the combination?

ELECTRICITY AND MAGNETISM

But in no case . . . is there a pure creation or production of power without a corresponding exhaustion of something to supply it.

—Michael Faraday, 1840

CHAPTER XXIV

ELECTROSTATICS

329. Historical. The fact was recorded by Thales of Miletus (640–546 B.C.) that when amber was rubbed with fur it acquired the ability to attract light bits of material, such as chaff, feathers, and hair. Other substances (hard rubber, sealing wax, glass, sulphur, etc.) were found when rubbed to possess the same property as amber.

TABLE OF CONDUCTIVITY (From Good Conductors to Poor Conductors)
Silver
Copper
Other metals
Carbon
Solutions of acids and salts
Animal tissue
Linen
Cotton
Dry wood
Paper
Silk
Glass
Mica
Amber
Sulphur
Hard rubber
Air

A body that has this property of attracting light bits of paper, silk, gold leaf, etc., is said to be **electrified,** from *electron*, the Greek word for amber; and its condition is due to what Sir Thomas Browne, in 1646, called "a charge of electricity." **Any quantity of electricity** is still referred to as **a charge.**

For many years certain materials, notably the metals, could not be electrified by rubbing. But Stephen Gray's discovery (1730) that electricity moves along, or is conducted by, some substances, suggested to Du Fay (1698–1739) that metals too could be electrified by rubbing if they were supported on glass, hard rubber, or sulphur. This was found to be true and led to the following classification of bodies:

Conductors, if they permit electricity to flow along them so readily that the effects of electrification are not observed; and

Nonconductors, or insulators, if they retain their charges of

electricity for a considerable time and hence exhibit the effects of electrification.

This distinction between conductors and nonconductors is not a sharp one. All substances can be arranged in a table according to their conductivities, as in the table above, from the best conductor (silver) at the top to the poorest conductor, or best insulator (dry air), at the bottom. But one cannot draw a line in this table and say that all above this line are conductors and all below it are insulators. The conductivity merely decreases gradually from the top to the bottom of the table. **There is no perfect conductor and no perfect insulator.**

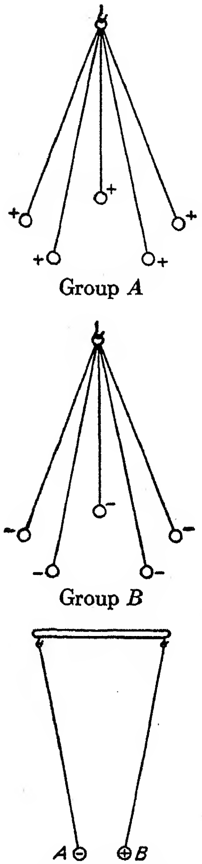


FIG. 278. Du Fay's Experiment

330. Du Fay's law. The first law of electrostatics was given in 1733 by Du Fay, after many experiments which may be summarized as follows:

If a group *A* of pith balls (Fig. 278), suspended by silk threads, is charged by contact with a piece of sealing wax that has been rubbed with fur; and if a second group *B* of pith balls similarly suspended is charged by contact with a glass rod that has been rubbed with silk, it is found that:

Any two balls of group *A* repel each other.

Any two balls of group *B* repel each other.

Any ball of group *A* attracts any ball of group *B*.

Any ball of group *B* attracts any ball of group *A*.

From such observations Du Fay concluded that there were two kinds of electricity: one developed on glass when rubbed with silk ("vitreous"); the other developed on sealing wax when rubbed with fur ("resinous"). He then gave the law as follows: **Like charges repel each other; unlike charges attract each other.** This law is entirely qualitative.

Nowadays we call the kind of electricity exhibited by glass when

rubbed with silk, **positive** (+); and that exhibited by sealing wax when rubbed with fur, **negative** (−), as suggested by Benjamin Franklin. But it should be carefully noted that these are just names for the two kinds of electricity, and do not always carry the same connotation as the terms "positive" and "negative" in algebra and geometry.

Conductors can be electrified as well as nonconductors, if they are supported by a nonconductor; and **any two different substances,** if held by insulating handles, **are electrified on being separated from contact.** The rubbing is not necessary; it serves only to secure a larger area of contact.

Electrostatic Series (From Positive to Negative)
+
Glass (with clean, polished surface)
Fur
Wool
Ivory
Quartz
Glass (passed through Bunsen flame)
Cotton
Silk
The Hand
Wood
Metals
Hard Rubber
Sealing Wax
Rosin
Sulphur
Gutta Percha
−

These facts enable us to test various kinds of material and arrange them in what is called an **electrostatic series,** such that if any two substances in the series are separated from contact, the one higher up in the series will be charged positively, and the lower one, negatively.

331. Electrical nature of matter. There is abundant evidence, which we shall see as we proceed, that matter is made up of four fundamental constituents: *

1. **Electrons,** the smallest particles of negative electricity, the mass of each being 1/1838 that of a hydrogen atom.
2. **Positrons,** the smallest particles of positive electricity, the mass and charge being equal to those of the electron in magnitude.
3. **Protons,** positive particles having a charge equal to that of a positron, but a mass 1837/1838 that of a hydrogen atom.
4. **Neutrons,** neutral particles (i.e., having no charge) with a mass equal to that of a proton.

* Theory seems to require the existence of a fifth particle, the "neutrino," having no charge and a mass equal to that of an electron. However, its existence has not yet been confirmed experimentally in a satisfying way.

The hydrogen atom consists of a proton as a central nucleus and one electron external to the nucleus.

All other atoms appear to consist of a nucleus, composed of protons and neutrons, and enough electrons external to the nucleus to make the atom as a whole neutral. The number of external electrons is therefore equal to the number of protons in the nucleus of a neutral atom.

This number is called the **atomic number** of the atom, and determines the position of the corresponding element in the periodic table.

Positrons, but recently discovered, appear to take no part in conduction.

Electrons, on the other hand, are abundant, are readily disengaged from the atom, and are much less massive than protons. Consequently, **when electricity is said to move along a conductor, the motion is of the electrons only,** except in the case of fluid conductors.

This theory supplies us with a very simple explanation of why any two bodies exhibit charges when separated from contact. For it is hardly reasonable to suppose that the electrons of two elements would be held to the nucleus with exactly the same force; consequently, we should expect more electrons to be disengaged from one kind of matter than from some other kind when they are separated after contact. The piece that loses electrons should exhibit a positive charge; and the other piece should exhibit an equal negative charge, for it has gained the electrons lost by the first piece of matter. The charges are indeed equal, as will be demonstrated in Sec. 334.

On the basis of this theory, we are able to account also for the difference between conductors and nonconductors. In conductors the electrons are rather loosely held to the nucleus, and readily move from molecule to molecule; whereas in nonconductors the electrons are firmly held by the nucleus and are far less free to drift from molecule to molecule. In solid conductors, the protons do not flow but remain in the atom of which they constitute part of the nucleus.

332. Electroscopes. An electroscope is a device for detecting the presence of a charge of electricity and its kind (+ or −). It

gives also a rough indication of the quantity of electricity constituting the charge.

A small ball of corn or alder pith suspended by a silk thread was one of the earliest forms of electroscope and was used by Henry Cavendish for quantitative determinations in 1773. This **pith ball electroscope** (Fig. 279) is greatly improved if the ball is coated with a conductor, such as aluminum or gold foil.

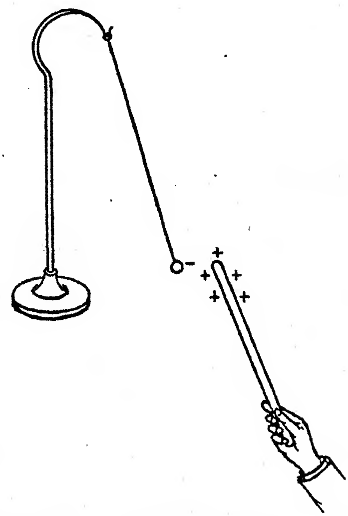


FIG. 279. Pith Ball Electroscope

Much more sensitive, however, is the **gold leaf electroscope** (Fig. 280), perfected by Abraham Bennet in 1786. It consists of two narrow strips of gold leaf *G*, *G*, each about 0.00001 in. thick, attached to a conducting rod *R* which projects through the insulating stopper *S* of the containing vessel *V*. The function of *V* is to prevent disturbance of the leaves by air currents. The rod usually terminates in a ball or in a flat plate (condensing electroscope). A small quantity of P_2O_5, or other drying agent, maintains a dry atmosphere within the vessel.

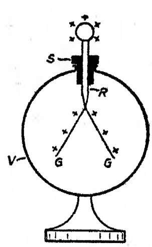


FIG. 280. Gold Leaf Electroscope

A part of any charge, applied to the conducting rod, distributes itself over the two gold leaves, which, being thereby charged with **like kinds** of electricity, repel each other, the angle of divergence being an index of the magnitude of the charge.

Refined forms of this electroscope have been used by the Curies for the measurement of radioactivity and by Millikan and Compton in the investigation of cosmic rays.

Coulomb's torsion balance (Fig. 290) is also in effect an electroscope; but **being provided with a scale, it is more properly called an electrometer.**

333. Methods of charging. When an excess of positive or negative electricity has been obtained on a body by rubbing or by an

influence machine (Sec. 369), we may charge another body or an electroscope by either of two methods:

1. *By contact*, when the charged body is touched against the body to be charged and either gives up electrons to it (negative charging), or takes electrons from it (positive charging).

2. *By induction.* In this method, discovered in 1729 by Stephen Gray, the charging body **does not come into contact** with the body to be charged. It is best described by means of the series of diagrams in Fig. 281. Let B_1 be a conductor on an insulating stand.

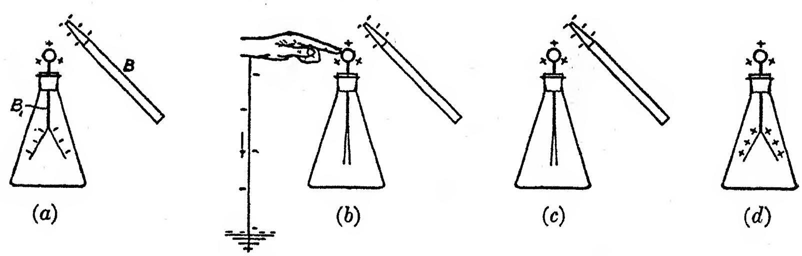


FIG. 281. Charging by Induction

In Fig. 281a, the charged body B, say, a piece of sealing wax that has been rubbed with fur, is brought near the body B_1 which is to be charged, but not in contact with it. In accordance with Du Fay's law, electrons are repelled to the parts of B_1 more remote from B, leaving the nearer parts with an excess of protons.

If B_1 is now touched with the finger (Fig. 281b), or otherwise connected to the earth by a conductor, the electrons will flow through this connection to the earth, leaving an excess of positive electricity on B_1. This positive charge does not escape because it is held by the attraction of the negative charge, and hence it is called the "bound charge."

Still holding the charging body near B_1, the earth connection is broken (Fig. 281c).

The charging body may then be removed, and B_1 is left with the excess positive charge (Fig. 281d).

It will be noted that in charging by contact, the charge acquired by the body to be charged is of the **same kind** as that on the charging body; whereas by induction, these charges are of **opposite kinds.**

334. Equality of charges.

(*a*) *Of charges obtained by separation from contact.* Let the fur with which a piece of ebonite is rubbed be made in the form of a close-fitting cap having a silk thread attached at the top (Fig. 282). If the cap is given several turns on the rod and is then with-

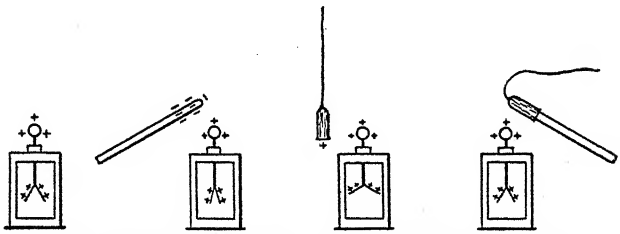


FIG. 282. Equal (+) and (−) Charges Are Developed

drawn by means of the thread, the ebonite will show a **negative** charge as usual when brought near a positively charged electroscope, for it will cause the leaves of the electroscope to diverge less.

The cap will likewise show that it possesses a charge, but **positive**, for it will cause the leaves of the electroscope to diverge more.

If the cap is then carefully replaced on the ebonite rod (being handled only with insulators), and the two (i.e., the rod and the cap) are brought near the ball of the electroscope, the leaves of the latter will be undisturbed, showing that the charges were equal, for they have neutralized each other.

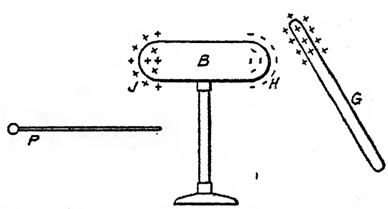


FIG. 283. Equal (+) and (−) Charges on Insulated Body

(*b*) *Of charges induced on an insulated conductor.* Let *B* (Fig. 283) be an insulated conductor. Hold near it the charged glass rod *G*. By means of a proof plane *P* (a small metal disk with an insulating handle), take a sample of the induced charge from the end *H* and test its sign at the electroscope. It will be found **negative.**

Similarly test the charge on the opposite end *J*. It will be found **positive.**

Remove *G* and again test *B* for charge at various points with

the proof plane. It will be found everywhere neutral. Hence the charges induced at H and J must have been equal, because when permitted to flow together they neutralized each other.

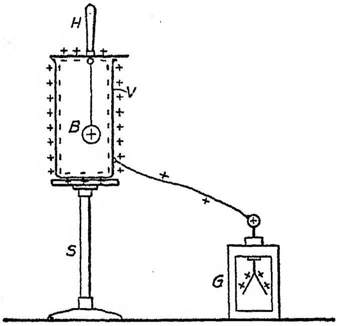


Fig. 284. Faraday's Ice-pail Experiment

(*c*) *Of the inducing and the induced charge* (Faraday's Ice-pail Experiment). Let a metal vessel V, or other hollow conductor (Fig. 284), rest on an insulating stand S and have its outer surface connected by a conductor to the knob of an electroscope G. A metal ball B is suspended by an insulating thread from the metal cover C. Let the cover be held by its insulating handle H and the ball charged positively, say.

Then if the ball is lowered into the hollow conductor, the leaves of the electroscope will diverge until the cavity is completely closed. Thereafter the ball may be moved about in the hollow conductor without causing any change in the divergence of G.

If the surfaces of V are tested with a proof plane just before it is closed by the cover, the inner surface will be found charged negatively and the outer, positively; for electrons have been drawn to the inside, leaving an excess of protons on the outside.

If the ball is removed, the vessel will test neutral as at first. Hence the **charges induced on its inner and outer surfaces were equal,** as in (*b*) above.

Let the ball again be placed within the vessel, and the leaves of G will diverge as before. If it is touched to any part of the inner surface, there will be no change in the divergence of the leaves, from which we may conclude that the charge on the outer surface has suffered no change. But on removing the ball, it will be found to be completely discharged, and the proof plane will show that there is no longer any charge on the inner surface. Since there was no change in the charge on the outer surface, **the charge on the inner surface must have been exactly equal and opposite to that on the ball,** and they consequently neutralized each other. Also, since the charges on the inner and outer surfaces were equal, **the charge on the ball must have been equal to the outer charge.**

Since the charging body was completely enclosed in the hollow conductor, we have here dealt with the total induced charges, just as we dealt with total charges in cases (*a*) and (*b*), above.

Thus, in all cases, the *total* positive and negative charges are equal, whether obtained by separation from contact or by induction.

335. Conservation of electricity. Since it appears from the preceding section that electrification consists not in making, or producing, electricity but merely in **separating the equal charges of opposite kinds** that were already present in the neutral body, it follows that **we can neither create nor destroy electricity**; or, the total quantity of electricity in the universe remains constant.

This statement requires a slight qualification. Recent research (see Sec. 657) has revealed that sometimes when a photon, or quantum of energy, disappears, an electron and positron are formed; and it appears probable that an electron and a positron will recombine to form a quantum of energy. But these transformations involve such minute quantities of electricity and energy that for practical purposes we may consider electricity to be conserved, just as are matter and energy.

336. Static charge on outer surface of a conductor. In Michael Faraday's ice-pail experiment it was observed that after the ball had been brought into contact with the inner surface, there was no charge on the ball or on the inner surface; but the charge on the outer surface remained as at first. It did **not** redistribute itself over both surfaces. Faraday demonstrated this fact by the following ingenious experiment.

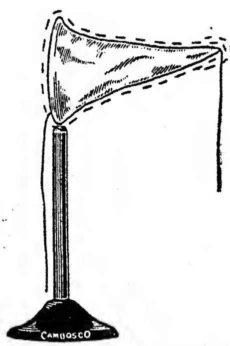


FIG. 285. Charge Resides on Outer Surface

A conical linen * bag (Fig. 285) is fastened at its large end to an insulated metal ring. It has a silk thread passing through it and extending about a foot beyond its apex and its base. This silk thread is fastened to the linen bag at its apex.

If the bag is charged with electricity, the proof plane and electroscope will show that **there is charge on its outer surface but none on the inner surface.**

* Linen is used because it is a fairly good conductor.

Let the silk thread then be drawn so as to turn the bag inside out, and its surfaces again tested with the proof plane. **The charge will be found now upon the new outer surface,** and **no charge will be found upon the inner surface,** which was originally the outer.

This result might have been predicted. The particles of electricity, being all alike, will repel each other. They consequently move through the conductor until each is in equilibrium as far from its neighbors as it can get. Hence we would expect to find them on the outside of the conductor, as we do.

337. Interior of hollow conductor shielded from external electrical disturbances (Faraday's Chamber Experiment). As a further test of the conditions within a hollow conductor, Faraday built a cubical box (each edge 12 ft long) and covered it with tin foil. This tin-foil chamber could be charged to a very high potential by an external electrical machine. He writes: "I went into the cube and lived in it, using lighted candles, electrometers, and all other tests of electrical states. I could not find the least influence upon them or indication of anything particular given by them, though all the time the outside of the cube was powerfully charged and large sparks and brushes were darting from every part of its outer surface."

The experiment may be repeated on a small scale by placing an electroscope inside a box of wire gauze, as in Fig. 286. When the gauze box is charged as highly as possible by an external influence machine M (Sec. 369), the electroscope is unaffected.

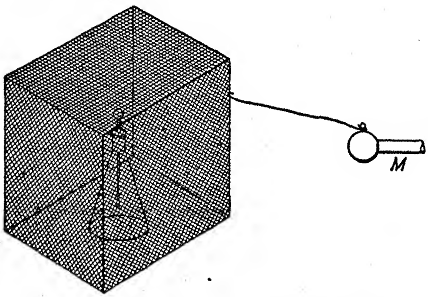


Fig. 286. Faraday's Chamber Experiment

This indicates also that a region may be shielded from external electrical fields by a metallic network as well as by a continuous metal housing. It follows that the safest place in an electrical storm is within a building completely covered with metal; but metal-roofed and metal-framed buildings are also very good protection.

338. Distribution of charge on an insulated conductor. Let a pear-shaped conductor B (Fig. 287) be mounted on an insulating stand and charged negatively, say. If we test the charge at various places by touching the proof plane flat against the surface of the conductor, we find that the leaves of the electroscope to which we apply the proof plane diverge much more widely when the charge is taken from the small end than when taken from the large end.

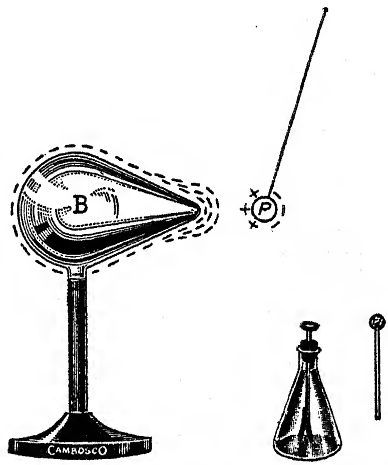


FIG. 287. Distribution of Charge

We conclude, therefore, that **the charge is most concentrated** (i.e., the surface density, or charge per unit area of the surface, is greatest) **where the curvature** * **is greatest** (radius least).

339. Effect of points. Let a pith ball covered with tin foil and suspended by a silk thread be brought near the small end of B (Fig. 287). Electrons of the coating of the ball are repelled to its farther side, leaving the nearer side positively charged, as shown. The ball is accordingly attracted to B until it touches. Thereupon, electrons from B flow on to the ball and neutralize its positive charge. Having then only a negative charge, it is repelled from B.

The curvature of sharp points approaches infinity. Hence there is very great concentration of electricity at points, and particles of matter in the air in the neighborhood of charged points are acted upon just like the pith ball of the preceding paragraph. The electricity upon them is separated into equal positive and negative charges by induction, and they are drawn to the points. On coming into contact with a point, the nearer charge, which is of opposite sign to that of the point, is neutralized; and the particle is left charged with the same kind of electricity as the point. Consequently it is repelled, as was the pith ball.

These repelled particles may be sufficient in number and speed

* In mathematics, curvature c is defined by $c \equiv 1/r$, where r is the radius of curvature.

to produce an "electric wind" strong enough to deflect the flame of a candle. If the charge concentrated on a point is great enough, electrons are disengaged from molecules of the adjacent air and these "ionized" molecules, or *ions*, are likewise repelled. When such ionization takes place, tiny sparks forming a "brush discharge" are visible at the points.

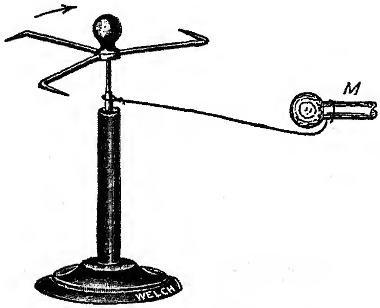


FIG. 288. Electricity Is Discharged by Points

The effect of points is well illustrated by the apparatus of Fig. 288.

A metallic wheel, having several spokes, sharp-pointed and bent as shown, is pivoted on a needle supported on an insulating base. When connected to a terminal of an electrostatic generator, the reactions of the repelled particles cause the wheel to rotate like a reaction turbine.

340. Protection from lightning. Benjamin Franklin (1706–90) was aware of this ability of points to discharge bodies, and applied it in his invention of the lightning rod (1750), which he insisted must be sharp pointed. These rods act in two ways: they tend to discharge the building and adjacent ground or the passing cloud; and they offer an easy and harmless path for the electricity in the case of a direct hit.

It is now well recognized that a building properly equipped with lightning rods is very effectually protected against serious lightning injury. The rods should be:

1. Sharp pointed at the top.
2. Extended several feet above the highest part of building.
3. Connected to all external metal work.
4. Not insulated from the building.
5. Well grounded in moist earth.

All high monuments, stacks, and buildings of modern construction are protected by lightning rods. On some buildings the rods show evidence of having been struck many times. The injury by lightning to valuable trees at the Johns Hopkins University has been greatly reduced by equipping them with lightning rods.

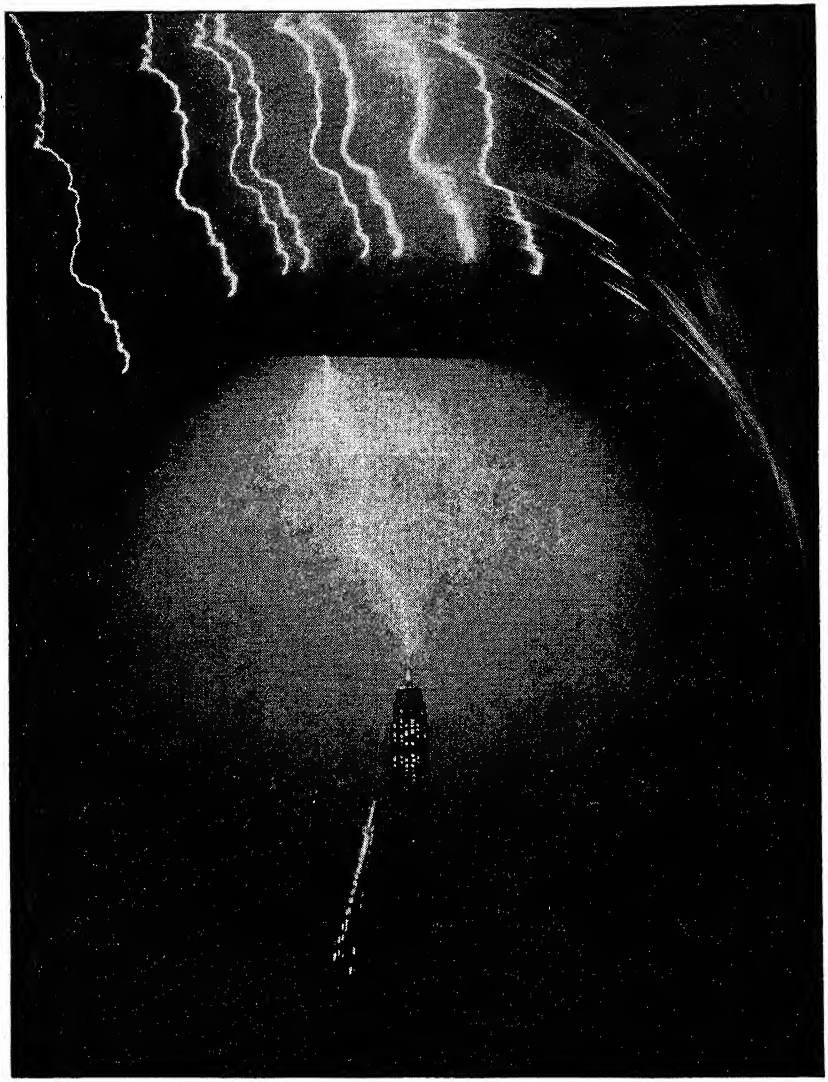


FIG. 289. Electricity Strikes Many Times in the Same Place. (Courtesy General Electric Co.)

In a lightning storm the safest places are:

1. An all-metal or steel-frame building.
2. A well-rodded building.
3. An automobile with metal top.
4. A cave, gulch, or foot of a high cliff.
5. Dense woods.

During such a storm one should:

1. Stay dry and indoors if possible.
2. Keep away from isolated trees, sheds, and barns, wire fences, open windows, and large bodies of metal, such as stoves.

341. Coulomb's law. The quantitative law obeyed by the force between two charges of electricity was evolved by Charles Coulomb, about 1787, by means of the torsion balance (Fig. 290), which he had devised in 1777. In substance, his observations were as follows.

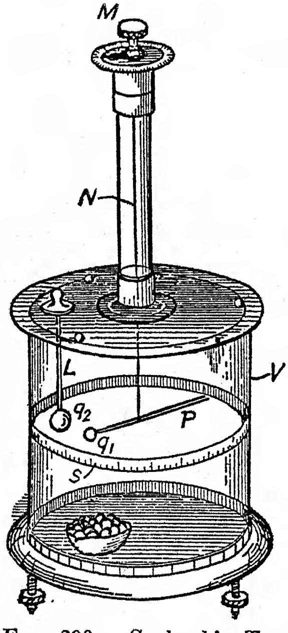


FIG. 290. Coulomb's Torsion Balance

A small, metal-coated pith ball is carried at one end of a pointer P, which is suspended by a very fine wire N from the torsion head M. Another ball attached to a glass rod L may be lowered through a hole in the cover of the cylinder. The distance between the centers of the two balls may be determined from the position of P as read on the scale S.

A charge q_1 is placed on the movable ball, and then various charges on the stationary ball. The force f of attraction or repulsion between these charges may then be determined from the amount that the wire must be twisted to bring the balls to a definite distance apart, this twist being measured by the angle through which M is turned.

Let q_1 and q_2 be the quantities of electricity on the balls of the balance;

r be the distance between the centers of the balls;

f be the force of attraction or repulsion between the two charges; and

k be the dielectric constant of the medium, such as air, between the charges.

Then it is found that:

$$f \propto q_1, \text{ when } q_2, r, \text{ and } k \text{ are kept constant}$$
$$\propto q_2, \text{ when } q_1, r, \text{ and } k \text{ are kept constant}$$
$$\propto \frac{1}{r^2}, \text{ when } q_1, q_2, \text{ and } k \text{ are kept constant}$$
$$\propto \frac{1}{k}, \text{ when } q_1, q_2, \text{ and } r \text{ are kept constant.}$$

Hence,

$$f \propto \frac{q_1 q_2}{kr^2} \quad \text{when all the factors vary}$$

and therefore,

$$f = C' \frac{q_1 q_2}{kr^2} \quad \text{for any units whatever} \tag{255}$$

where C' is a constant depending upon the **choice** of units.

Equation (255) is **Coulomb's law** in algebraic form.

Expressed in words, Coulomb's law states that **the force f of attraction or repulsion between two point charges q_1 and q_2 varies directly as their product and inversely as the dielectric constant k of the medium and the square of the distance r between them; but only when they are in a homogeneous isotropic medium.**

The direction of the force on each charge is along the line joining the charges. If the charges are both positive or both negative, f will be + and the charges will repel each other, according to Du Fay's law. If one is + and the other −, f will be − and they will attract each other.

It should be noted that the law is true only for **point charges,** i.e., when each charge may be considered concentrated at a point, and for uniformly charged spheres whose radii are small compared to the distance r between their centers. The law does not apply to bodies in general.

The effect of the medium as represented by the dielectric constant k was discovered by Faraday some 50 years after Coulomb's

work. The method of determining the value of k for a given medium is discussed in Sec. 364.

342. Electrostatic units. Up to this point no unit of charge, or quantity, of electricity has been chosen. Coulomb's law was established by using arbitrary quantities, which he obtained as follows. The stationary ball was first given any charge q_2, and the force on any other charge q_1 was determined. The sphere carrying q_2 was then removed and touched to another metal-covered sphere of exactly the same size. Since the two spheres were exactly alike, he assumed that the charge divided equally between them, so that on separation the first sphere had half its original charge. It was then replaced in the torsion balance, and the force on q_1 at the same distance r was measured. Similarly, by touching the first sphere to two conducting spheres exactly like it, there remained on it only one-third of its original charge, and so on.

Having obtained the algebraic form of Coulomb's law, we are now in a position to adopt a scientific system of units for electrostatics. For Eq. (255) it will be observed that we have already an absolute unit of force f (the dyne) and an absolute unit of distance r (the centimeter), but **no unit** for the q's and **none** for k. Units for these last two quantities must be defined.

As already mentioned, the constant of proportionality C' of Eq. (255) depends upon the units chosen; consequently we may choose our units so as to make it any number we please. The most convenient value C' could have would be unity, for that would simplify the equation. We therefore proceed to select a system of units that will make $C' = 1$.

First we define **the unit of dielectric constant as the dielectric constant of a vacuum.** That is, $k \equiv 1$ for a vacuum. Then we define: **The statcoulomb (stc), or electrostatic unit quantity, of electricity is that quantity which will repel an equal like quantity at a distance of one centimeter in a vacuum with a force of one dyne.**

Substituting the above values in Eq. (255), we have

$$1_{\text{dyne}} = C' \frac{1_{\text{statcoulomb}} \times 1_{\text{statcoulomb}}}{1_{\text{for vacuum}} \times (1_{\text{cm}})^2}.$$

Solving this for C', $C' = 1$, which was what we set out to secure.

Since C' is equal unity for these *absolute electrostatic units* (esu), we may write Coulomb's law in the simplified form:

$$f = \frac{q_1 q_2}{kr^2} \quad \text{provided esu are used.} \tag{256}$$

By his oil-drop experiment (see Sec. 373), Millikan has shown that 1 electron = 4.770×10^{-10} statcoulomb.* This appears to be the smallest package of electricity that occurs free in nature. Here the word "electron" is used as a unit of quantity, which was the sense in which its use was originally proposed by G. Johnstone Stoney in 1891.

In practical work a much larger unit quantity of electricity is desirable, for which we use the *coulomb*. This is defined in Sec. 396; but for present purposes, as a good approximation:

$$\begin{aligned} 1 \text{ coulomb} &= 3 \times 10^9 \text{ statcoulombs} \\ &= 6.3 \times 10^{18} \text{ electrons.} \end{aligned}$$

343. The electric field. The region about an electric charge where its influence may be detected is called its *electric field.*

The **test charge** by means of which an electric field is investigated is taken to be a **unit positive charge.** Actually a unit charge is seldom used, but results are reduced to what they would have been had a unit charge been used.

The **intensity,** or strength, of an electric field at a point A is defined as the force in dynes which acts upon a unit positive charge when placed at the point A.

Algebraically,

$$\mathfrak{F} = \frac{f}{q_1} \; \frac{\text{dynes}}{\text{stc}} \tag{257}$$

where $\mathfrak{F}$ is the intensity in the field at a certain point in dynes per statcoulomb;
q_1 is the (+) charge in statcoulombs placed at the point in question; and
f is the total force in dynes that acts on q_2.

Field intensity, being a vector divided by a scalar, is a vector.

The **direction** of an electric field at a point A is defined as the direction of the field intensity at the point A.

* More recent investigator finds 1 electron = 4.802×10^{-10} stc.

344. A line of force is the path described by a + point charge moving always in the direction of the field intensity. Thus, if a +1 charge is placed in the neighborhood of two small spheres uniformly charged with $+q$ and $-q$ stc, respectively, it will be

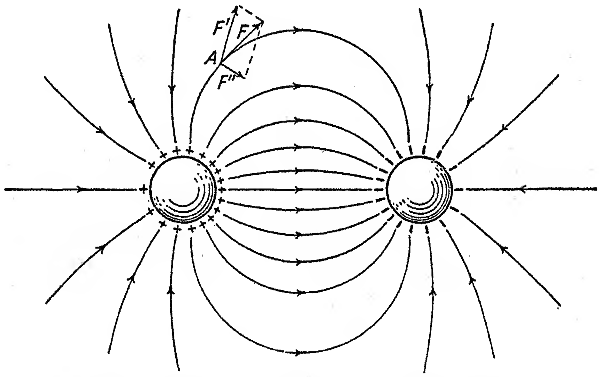


FIG. 291. Lines of Force between Unlike Charges

acted upon by the field intensity F' due to $+q$ and by F'' due to $-q$. Their resultant F will be the total field intensity at the point A, and will be tangent at A to the line of force that the +1 charge would follow. (See Fig. 291.)

Such lines of force may be mapped by using a small paper "needle" suspended on a single fiber of unspun silk. They are more easily obtained by scattering short bits of unraveled silk thread, or clippings from a camel's hair brush, over the region near the charges. Unlike charges will be induced on the opposite ends of the bits of silk, which will then arrange themselves along the lines of force.

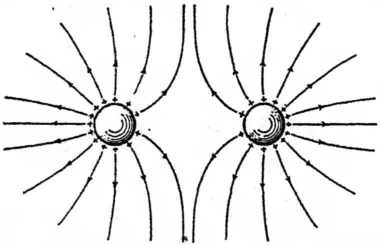


FIG. 292. Lines of Force between Like Charges

It will be observed that the lines of force bulge outward as if they repelled one another and that they do not cross. They start on a + charge and end on a − charge. The lines may be thought of as in tension, tending like elastic threads to draw the + and − charges together. But there is this difference: as elastic threads shorten the tension becomes less, whereas when lines of force shorten, the forces drawing the charges together become greater.

Figure 292 shows the lines of force in the neighborhood of two like charges.

We may use lines of force to represent the electric field intensity *quantitatively*. Thus, suppose at the point P (Fig. 293) the field intensity $\mathfrak{F}$ is 5 dynes/stc. This would be represented graphically by taking a plane surface 1 cm square, perpendicular to the direction of the field, and having P as its center, and by drawing (uniformly distributed through this area) 5 lines in the direction of the field. When an electric field is represented in this way, one can get at a glance a fair idea of its intensity at different points. Where the lines are close together it is strong; where they are widely separated, it is weak.

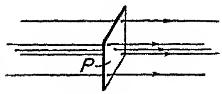


FIG. 293. Lines Used to Indicate Field Intensity

The concept of lines of force was introduced by Faraday, who used it very effectively to describe the behavior of electrified and magnetized bodies. However, it should be understood that lines of force are only a convenient means of representing fields of force graphically, and do not actually occur in nature.

345. Electric field intensity due to a point charge. Let a point charge of $+q$ stc be placed at some point P (Fig. 294) in a very large region of space filled with a medium of dielectric constant k, but otherwise empty. We wish to find the field intensity $\mathfrak{F}$ at a point A, whose distance from P is r cm, due to the charge $+q$.

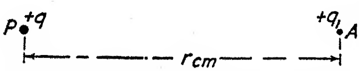


FIG. 294. Field of a Point Charge

Imagine another point charge of $+q_1$ stc placed at the point A. Then each charge repels the other with a force f whose value is given by Coulomb's law:

$$f = \frac{qq_1}{kr^2}.$$

But this is the force on $+q_1$ stc, and the field intensity is defined as the force on $+1$ stc. Therefore,

$$\mathfrak{F} \equiv \frac{f}{q_1} = \frac{\frac{qq_1}{kr^2}}{q_1} = \frac{q}{kr^2}\,\frac{\text{dynes}}{\text{stc}} \tag{258}$$

where q is the charge that produces the field.

346. Electric flux density, or displacement. Consider two parallel plane conductors A and B, normal to the paper, their length and breadth being large compared with the distance between them; and let them be kept uniformly charged by a generator G with $+q$ and $-q$ stc per cm², respectively, in a vacuum.

Between two such plates the field intensity ($\mathfrak{F}$ dynes/stc is uniform and normal to the plates, except near the edges. This may be shown with bits of unraveled silk thread or camel's hair, and is easily seen to be true theoretically after Sec. 357.

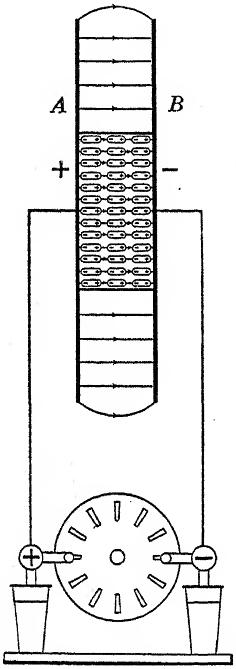


FIG. 295. Dielectric Flux

Let a rectangular prism of dielectric be fitted between the plates, with its lateral edges parallel to the lines of force, as shown in Fig. 295.

The field intensity within the dielectric will then be the same as in the vacuum between the plates. But the particles of the dielectric will be polarized; that is, the electrons will be displaced from their normal positions in the atoms in a direction opposite to that of the field. Hence, to represent the condition within the dielectric we should draw additional lines between these little pairs of + and − charges that are due to polarization; and for a given field intensity F there would be more lines than in the adjacent vacuum.

Since this increase of lines within the dielectric arises from the displacement of the atomic charges, the total number of lines per cm² is called the **electric displacement,** or, more commonly, the **electric flux density.** It turns out that if we define the flux density D to be k times the field intensity $\mathfrak{F}$, where k is the dielectric constant, the relations derived therefrom lead to results that are in accord with experiment. Hence,

$$D \equiv k\mathfrak{F}. \tag{259}$$

Like field intensity, therefore, flux density may be represented by lines drawn as described for field intensity, the number per

cm² being the numerical value of D instead of $\mathfrak{F}$. However, these lines should be called **lines of displacement,** not lines of force.

The **unit of flux density** is called **one line per square centimeter.** It is the flux density at a point in a medium of dielectric constant 1 when the field intensity at that point is 1 dyne per statcoulomb.

Since $\mathfrak{F}$ is a vector and k a scalar, D is a vector, having the same direction as $\mathfrak{F}$ in isotropic media.

We seldom actually draw the lines to represent either field intensity or flux density. But the concept of these lines, even though they have no actual existence, is most helpful in giving us a suggestive language in which to discuss electrostatic problems.

347. The **flux density at a point due to a point charge** is found at once by substituting in Eq. (259) the value of $\mathfrak{F}$ from Eq. (258):

$$D \equiv k\mathfrak{F} = k\frac{q}{kr^2} = \frac{q}{r^2}\ \frac{\text{lines}}{\text{cm}^2} \tag{260}$$

where D is the flux density at the point in question;
q is the charge in stc that produces the field; and
r is the distance in cm from q to the point in question.

Here it will be noticed that k has disappeared from the value of D. Hence the flux density produced by a given charge is independent of the medium; and lines representing flux density are continuous through different media, although in general they are bent at the surface of separation of two media unless normal thereto.

348. Total electric flux. The total electric flux Ψ is defined as the total number of lines required to represent the entire field under consideration by the method described in the preceding paragraph.

If A is the total area through which the flux passes perpendicularly, and if the flux density D is *uniform,* then

$$\Psi \equiv DA. \tag{261}$$

In case the flux density is not uniform throughout the whole area, we should divide the area up into little areas dA, so small that the flux through each could be considered uniform. The definition then becomes for each dA:

$$d\Psi \equiv D\,dA. \tag{262}$$

349. Total flux from a charge q: Gauss' law. The effect of a charge uniformly distributed over a sphere is the same outside the sphere as if it were all concentrated at the center. Let a charge of $+q$ stc be concentrated at a point or uniformly distributed on a small sphere. We wish to compute the total flux associated with the charge $+q$.

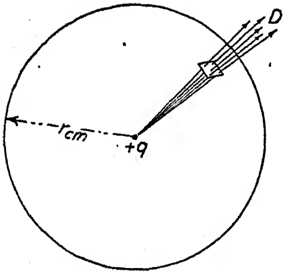


FIG. 296

About the charge, let a sphere of radius r cm (Fig. 296) be described, having its center at the center of the charge.

At any point on the spherical surface, the field intensity is:

$$\mathfrak{F} = \frac{q}{kr^2} \frac{\text{dynes}}{\text{stc}} \qquad \text{radially outward;}$$

hence the flux density at any point is

$$D = \frac{q}{r^2} \frac{\text{lines}}{\text{cm}^2} \qquad \text{radially outward.}$$

Since this is true for any cm² taken at random, it is true for every cm² of the sphere. From geometry, the area of the sphere is:

$$A = 4\pi r^2 \text{ cm}^2.$$

Hence the total flux from the charge $+q$ is:

$$\Psi \equiv DA = \frac{q}{r^2}(4\pi r^2)$$

$$= 4\pi q \text{ lines,} \qquad \text{radially outward.} \qquad (263)$$

This relation is known as **Gauss' law** for electric charges.

350. Electric potential at a point. Consider a region of space so large as to be sensibly free from external influences and filled with a medium of dielectric constant k, but otherwise empty. Let a small charge $+q_1$ be placed anywhere in the region. It may be moved about at will without any work being done, because there is no force to be overcome since there is no other charge in the region to exert force upon it.

But let a charge $+q$ be first placed at any point P in this region,

and the small charge $+q_1$ then brought in. To move $+q_1$ from any point J (Fig. 297) to any other point G **will now require work,** for $+q_1$ will be acted upon by the repulsive force from $+q$, which must be overcome along any path from J to G. Hence the charge

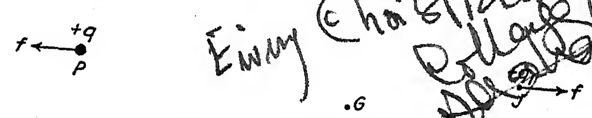


FIG. 297. The Idea of Potential.

$+q$ confers upon the region at every point a new property, whereby work must be done to move any other charge up to that point. This property is called **potential.**

Qualitatively, **electric potential is that property of a system at a point in virtue of which work must be done to bring a charge of electricity up to that point.**

Electric potential V at a point is measured by the amount of work required to bring a $+1$ charge from infinity up to that point. Hence, in general, if W units of work are required to bring Q units of electricity from infinity up to a certain point, the potential V at that point is defined by the equation:

$$V \equiv \frac{W}{Q}. \tag{264}$$

Potential is a scalar quantity, since both W and Q are scalar quantities.

351. Units of potential. The electrostatic unit of potential, the *statvolt*, is obtained by putting unity for each quantity in Eq. (264):

$$1_{\text{statvolt}} \equiv \frac{1_{\text{erg}}}{1_{\text{statcoulomb}}}. \tag{265}$$

This may be stated in words as follows:

The statvolt (stv) is defined as the potential at a point when one erg of work is required to bring one statcoulomb of electricity from infinity to the point in question.

While this is the most convenient unit of potential for electrostatics, it is rather large for practical purposes. The **practical unit** of potential is called the volt.

The **volt** is the potential at a point when one *joule* of work is required to bring one *coulomb* of electricity from infinity to the point in question. That is,

$$1 \text{ volt} \equiv \frac{1 \text{ joule}}{1 \text{ coulomb}}. \tag{266}$$

Since 1 joule $\equiv 10^7$ ergs, and 1 coulomb $= 3 \times 10^9$ statcoulombs, we have:

$$\begin{aligned} 1 \text{ volt} &= \frac{10^7 \text{ ergs}}{3 \times 10^9 \text{ statcoulombs}} \\ &= \frac{1}{300}\left(\frac{1 \text{ erg}}{1 \text{ statcoulomb}}\right) \\ &= \frac{1}{300} \text{ statvolt.} \end{aligned} \tag{267}$$

Attention is called to the fact that Eq. (266) is a working relation, which may be considered a secondary definition. The primary definition of the volt is given in Sec. 403.

352. Potential due to a point charge. In Fig. 298, let $+Q$ be a point charge at P in a medium whose dielectric constant is k, there being no other charges in or near the region.

By Eq. (258), $\mathfrak{F} = 0$ when $r = \infty$; and as the $+1$ test charge gets nearer to the $+Q$ charge, $\mathfrak{F}$ increases but not linearly. It is necessary, therefore, to add up the amounts of work done in infinitesimal steps dr, taken so small that in each step $\mathfrak{F}$ may be considered constant.

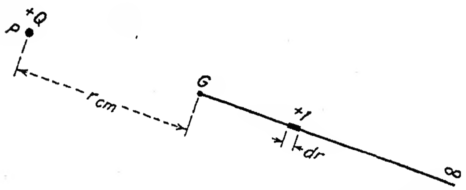


Fig. 298. Potential Due to a Point Charge

When the $+1$ charge is at any distance r from P, it is acted upon by the force $\mathfrak{F} = Q/kr^2$ dynes, radially away from P. To move the test charge toward P will therefore require a force at least as great as $-Q/kr^2$ dynes; and the work done in the distance dr will be $-(Q/kr^2)dr$ ergs.

The potential V at G is, by definition, the total work done on the system in bringing the $+1$ test charge from ∞ up to G, which is at the distance r cm from $+Q$. The potential is therefore the sum of these infinitesimal amounts of work.

Stating this in mathematical symbols,

$$V = \int_{\infty}^{r} - \frac{Q}{kr^2}dr = \left[\frac{Q}{kr}\right]_{\infty}^{r} = \frac{Q}{kr} \text{ ergs/statcoulomb.} \qquad (268)$$

In this case, work is done on the system by an external force in pushing the +1 charge up to G in opposition to the field intensity due to Q. The potential at G is positive, and the system gains energy with each +1 charge so brought up. Had the charge at P been $-Q$, the +1 test charge would have been attracted up to G against an external force; i.e., the system would have done work on an external force and would have lost energy. Or we may say that an external force did negative work on the system in bringing up the +1 charge, and the potential would accordingly be negative.

353. Potential difference. Since potential is a scalar quantity, the potential difference V between two points J and G is the algebraic difference between the potential V_2 at J and V_1 at G (Fig. 299).

$$V \equiv V_2 - V_1.$$

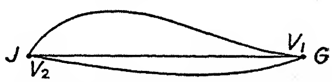


FIG. 299. Potential Difference Is Independent of Path

Obviously V* is a quantity of the same kind as V_2 and V_1; and, **potential difference will be measured in the same units as potential.**

Also, since V_2 and V_1 are numerically equal to the work required to bring a +1 charge from ∞ up to J and G, respectively, the potential difference V is numerically equal to the work required to carry a +1 charge from G to J.

That is, **the potential difference between two points is measured by the work required to carry a +1 charge from the first point to the second.** It is independent of the path followed in going from one point to the other.

Algebraically,

$$V \equiv \frac{W}{Q}. \qquad (268a)$$

where W is the work required to carry the charge Q from one point to the other.

* The use of V for both potential and potential difference causes no confusion, as the context will always make clear which is meant.

354. Effect of a uniformly charged sphere. A uniform charge is most easily secured on a sphere by using a conducting sphere supported by an insulator, say, a silk thread. A charge applied to such a sphere will spontaneously distribute itself uniformly over the surface of the sphere, on account of Du Fay's law and the symmetry of the figure.

Outside of such a charged sphere, the field intensity, flux density, and potential at any point may be shown to be the same as if the entire charge were concentrated at the center of the sphere.

Inside the sphere, the field intensity and flux density are each zero in consequence of the experiment of Sec. 337; and the potential at all interior points is the same as the potential at the surface.

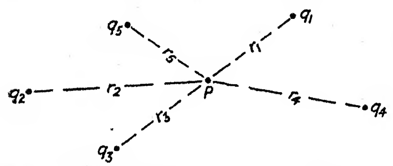


FIG. 300. Potential Due to a Number of Charges

355. Total potential at a point due to a number of charges. Since potential is a scalar quantity, the potential V at any point P (Fig. 300) due to the various charges q_1, q_2, q_3, etc., is the ordinary algebraic sum of the potentials at P due to each of the charges separately. That is,

$$V = \frac{q_1}{kr_1} + \frac{q_2}{kr_2} + \frac{q_3}{kr_3} + \ldots, \text{ or, more briefly,}$$

$$V = \Sigma \frac{q}{kr}. \tag{269}$$

356. Tendency of electricity to flow from places of higher to places of lower potential. In Sec. 79 it was seen that if friction is neglected, a body has potential energy equal to the work that was done in getting it into the position or condition that it has.

From its definition, potential at a point is equal (numerically) to the work required to bring a unit positive charge from infinity up to that point. Consequently, we should expect from Sec. 79 that when a +1 charge has been brought up to a point, it would have potential energy equal (numerically) to the potential at the point. And if this is true, we should expect electricity to conform to the principle of Sec. 80, that a system tends to adjust itself so that its potential energy is a minimum.

That is, **electricity should tend to flow from a place of higher to a place of lower potential**; and experiment shows that **it does.**

357. Equipotential surfaces. A surface all points of which are at the same potential is called an *equipotential surface.* The following facts with regard to an equipotential surface are important:

(a) **No work is done in moving a charge from point to point on an equipotential surface.** For if work were done in carrying a +1 charge, say, from J to G on equipotential surface $MGJN$ of Fig. 301, then G would be at a higher potential than J, which is contrary to the hypothesis that the surface is equipotential.

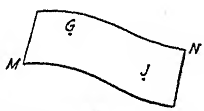


FIG. 301. Equipotential Surface

(b) **The field intensity F is always normal to an equipotential surface.** For if it were not, it would have a component F' parallel to the surface (Fig. 302), and this would cause electricity to flow along the surface. But that would indicate that P was at a higher potential than some neighboring point, which is contrary to the hypothesis that the surface is equipotential.

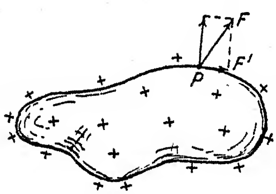


FIG. 302. Intensity is Normal to Equipotential Surface

Since the direction of the flux at a point is always the same as that of the field intensity, it follows that **the flux is everywhere normal to an equipotential surface.**

(c) **A conductor in which no electricity is flowing is an equipotential body.** For if it were not, electricity would flow through the conductor from points of higher to points of lower potential; and this is contrary to the hypothesis that no electricity is flowing. Hence a conductor having a static charge (which may be zero) has an equipotential surface and is equipotential throughout. If there is an empty cavity in the conductor, the potential at all points within the cavity is the same as that of the conductor. For, there being no charge in the cavity, there is no flux; and hence no field intensity and therefore no potential difference between points.

358. Earth's potential arbitrarily zero. For convenience, we need some body whose potential may be taken as zero, with reference to which the potentials of other bodies may be measured.

The earth, being on the whole a good conductor and accessible to observers everywhere, is the most suitable body for this purpose. Accordingly, *the potential of the earth is arbitrarily taken to be zero.*

359. Condensers. Let a sheet of metal M be mounted on an insulating base and connected by a wire to an electroscope G, as in Fig. 303; and let them be charged negatively until the leaves of G diverge widely.

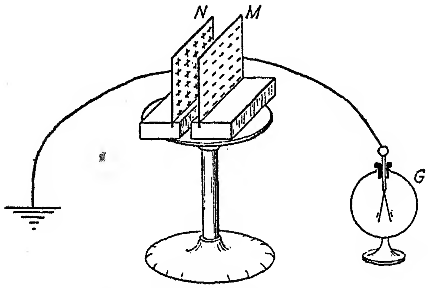
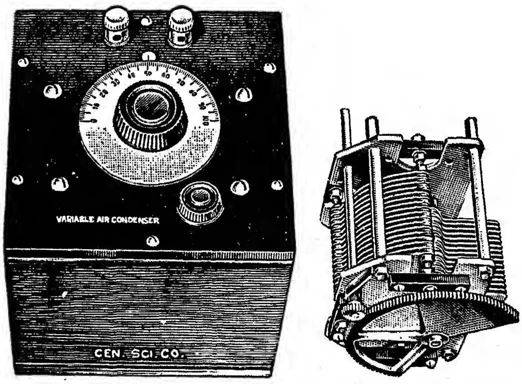


Fig. 303. Demonstration Condenser

If a second sheet of metal N, **connected with the earth,** is now brought near M, the leaves of the electroscope will close considerably. The negative charge on M repels electrons to the earth, leaving a + charge bound on N. This + charge on N attracts electrons to the nearer face of M, thereby reducing the number on the electroscope, so that the leaves fall.

In this action, the presence of the grounded plate N causes electrons to be more concentrated, or condensed, on plate M; hence the arrangement is called a "condenser." **An electric condenser consists of two conductors separated by an insulator.**

Parallel plate condensers, consisting of a number of metal plates mounted parallel and separated by sheets of glass, oil, or



(a) (b)

Fig. 304. Variable Air Condenser. (Courtesy Central Scientific Co.)

air, are used extensively. The familiar variable air condenser of radio sets is of this type, shown in Fig. 304. **Telephone condensers** consist of two sheets of tin foil 3 or 4 in. wide and from 10 to 20 ft long, separated by paraffin paper and rolled into a compact bundle. The complete condenser is encased in a rectangular metal can for protection (Fig. 305). This condenser has large capacitance, small size, and convenient shape.

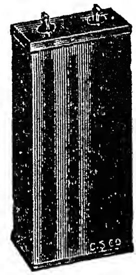


FIG. 305. Telephone Condenser. (Courtesy Central Scientific Co.)

360. Capacitance. Condensers are valuable because of a property now called "capacitance," but formerly called "capacity." A condenser of fixed capacitance is conventionally represented by Fig. 306; one of variable capacitance, by Fig. 307.

Consider a condenser D connected, as shown in Fig. 308, to a battery B and a ballistic galvanometer G (Sec. 400) for measuring quantity of electricity. Suppose the cells of B to be all alike, each developing a potential difference of V_1 units.

FIG. 306

Let the movable contact of B be set so that 1 cell only is in the circuit, and place the switch S on contact 1. The condenser then has a potential difference V_1 maintained between its terminals; and a certain quantity of electricity (electrons) will flow from its positive to its negative plate. This quantity, without regard to sign, is called *the charge on the condenser.*

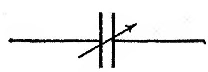


FIG. 307

If the switch is then suddenly thrown from contact 1 to contact 2, this charge on the condenser will be discharged through the ballistic galvanometer, which will indicate its quantity Q_1.

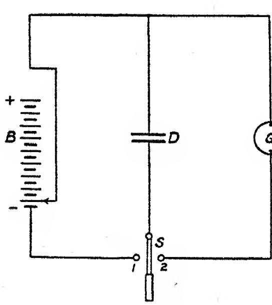


FIG. 308

Let the experiment be repeated, using two cells. This will double the potential difference between the plates of the condenser, and we might expect that double the quantity of electricity would be transferred from its + to its − plates. On discharging through G, we find the quan-

tity to be exactly $2Q_1$, as was anticipated. Similarly, using three cells, we get $3Q_1$, and so on.

Tabulating these values of the potential difference E between the plates and the charge Q obtained from the condenser, we have:

Values of V	Values of Q
V_1	Q_1
$2V_1$	$2Q_1$
$3V_1$	$3Q_1$
$\vdots$	$\vdots$

from which it is seen at once that Q is proportional to V.

Algebraically,

$$Q \propto V.$$

Therefore,

$$Q = CV. \tag{270}$$

The constant of proportionality C is called the *capacitance* of the condenser. Solving for C,

$$C \equiv \frac{Q}{V} \tag{271}$$

which is the defining equation of capacitance.

In words, **capacitance is the ratio of the charge on a condenser to the potential difference between its terminals.**

361. Units of capacitance. Units of capacitance are obtained by making each quantity unity in Eq. (271). In the electrostatic system,

$$1_{\text{statfarad}} \equiv \frac{1_{\text{statcoulomb}}}{1_{\text{statvolt}}}. \tag{272}$$

The unit of capacitance so defined is called the statfarad as indicated above. In words, **the statfarad is the capacitance of a condenser which, when charged with one statcoulomb of electricity, has a potential difference of one statvolt between its plates.**

In the practical system,

$$1_{\text{farad}} \equiv \frac{1_{\text{coulomb}}}{1_{\text{volt}}}. \tag{273}$$

That is, **the farad is the capacitance of a condenser which, when**

charged with one coulomb of electricity, has a potential difference of one volt between its plates.

Solved Problem

Compute the relation between a farad and a statfarad.
Known:

$$1 \text{ farad} \equiv \frac{1 \text{ coulomb}}{1 \text{ volt}} \qquad \text{(a)} \qquad \text{by Eq. (273)}$$

$$1 \text{ coulomb} = 3 \times 10^9 \text{ statcoulombs} \qquad \text{by Sec. 342}$$

$$1 \text{ volt} = \tfrac{1}{300} \text{ statvolt.} \qquad \text{by Eq. (267)}$$

Solution: Substituting these values for the coulomb and the volt in Eq. (a),

$$1 \text{ farad} = \frac{3 \times 10^9 \text{ statcoulombs}}{\frac{1}{300} \text{ statvolt}}$$

$$= 9 \times 10^{11} \frac{1 \text{ statcoulomb}}{1 \text{ statvolt}}.$$

But

$$\frac{1 \text{ statcoulomb}}{1 \text{ statvolt}} \equiv 1 \text{ statfarad.} \qquad \text{by Eq. (272)}$$

Therefore

$$1 \text{ farad} = 9 \times 10^{11} \text{ statfarads.} \qquad (274)$$

Although the farad is the recognized unit of capacitance in the practical system of units, it is far too large for most practical purposes. The unit of capacitance most commonly used is the **microfarad** (μf):

$$1 \text{ microfarad} \equiv 10^{-6} \text{ farad} \qquad (275)$$

$$= 9 \times 10^5 \text{ statfarads.}$$

362. Capacitance of an isolated sphere. Given an isolated sphere of radius r cm, uniformly charged with Q stc of electricity, it is required to find the capacitance of the sphere with reference to a plane at infinity.

By Sec. 354, a charge uniformly distributed over a sphere produces the same effects **outside** the sphere as if it were concentrated at its center. Hence, if a sphere of radius r cm is charged with Q stc of electricity, the potential of its surface due to its own charge is, by Eq. (268):

$$V = \frac{Q}{r}$$

if the medium surrounding it is a vacuum or air ($k = 1$).

This V is likewise the potential difference between the sphere and a body, say, a plane, at infinity, for it is the work required to bring a $+1$ charge from infinity to the sphere. Therefore,

$$C \equiv \frac{Q}{V} \qquad \text{by Eq. (271)}$$

$$= \frac{Q}{\left(\frac{Q}{r}\right)}$$

$$= r \quad \text{statfarads.} \tag{276}$$

That is, the capacitance of a sphere in statfarads, with reference to a plane at infinity, is numerically equal to its radius in centimeters.*

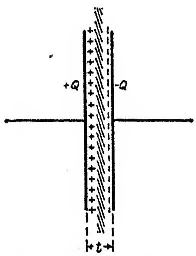


FIG. 309. Parallel Plate Condenser

363. Capacitance of a parallel plate condenser. In a given parallel plate condenser consisting of two conducting plates charged with $+Q$ and $-Q$ esu of electricity, respectively, and separated by a layer of dielectric whose thickness is t and dielectric constant k (Fig. 309), the total flux between the plates will be:

$$\Psi = 4\pi Q \text{ lines} \qquad \text{by Eq. (263).}$$

If we assume the charges to be uniformly distributed, this will give a flux density of

$$D = \frac{4\pi Q}{A} \frac{\text{lines}}{\text{cm}^2} \qquad \text{by Sec. 346}$$

normal to the plates, except at the edges, provided the dimensions of the plates are large compared to the distance t between them.

Hence, by Sec. 346

$$\mathfrak{F} = \frac{D}{k} = \frac{4\pi Q}{kA} \frac{\text{dynes}}{\text{stc}}.$$

The potential difference V between the plates is, by definition,

* This fact led to the use of the term "centimeter" as the esu of capacitance before the statfarad was adopted.

the work required to move a unit + charge from the negative to the positive plate, against the field intensity. That is,

$$V = Ft \cos 0^\circ \qquad \text{by Sec. 76}$$

$$= \frac{4\pi Q}{kA}t \text{ stv.}$$

Therefore,

$$C \equiv \frac{Q}{V} = \frac{Q}{\frac{4\pi Q}{kA}t}$$

or

$$C_1 = \frac{kA}{4\pi t} \text{ statfarads.} \tag{277}$$

This is the value of C for a condenser having two conducting plates (i.e., one layer of dielectric). If there are n conducting plates there will be $n - 1$ layers of dielectric, and the capacitance will be:

$$C_n = \frac{kA}{4\pi t}(n - 1) \text{ statfarads.} \tag{278}$$

364. Determination of dielectric constant. Equation (277) offers an obvious method for determining the dielectric constant of a substance. Let a parallel plate condenser be carefully constructed and placed in a vacuum. Its capacitance C_v will then be:

By Eq. (277)

$$C_v = \frac{k_v A}{4\pi t}. \tag{a}$$

Let the space between its plates be then filled with the substance whose specific inductive capacitance k_s is to be determined. The capacitance C_s will now be:

By Eq. (277)

$$C_s = \frac{k_s A}{4\pi t}. \tag{b}$$

Dividing Eq. (b) by Eq. (a),

$$\frac{C_s}{C_v} = \frac{k_s}{k_v}$$

and since $k_v = 1$ by definition,

$$k_s = \frac{C_s}{C_v}. \tag{279}$$

That is, the dielectric constant, or specific inductive capacitance, k_s of a substance is equal to the ratio of the capacitance of a condenser when that substance is used as the dielectric, to the capacitance of the same condenser when vacuum is the dielectric.

TABLE OF SPECIFIC INDUCTIVE CAPACITANCES

Vacuum	1.0	Paper	2.6
Hydrogen	1.00026	Wood (dry)	2 to 6
Air (dry)	1.00059	Mica	3 to 6
Paraffin	2.1	Bakelite	5
Hard Rubber	2.5	Glass	5 to 9
Oil	2.5	Water	81

365. Series and parallel connections. A number of conductors are said to be connected **in series** when all the electricity flows through each conductor. This type of connection is represented diagrammatically in Fig. 310.

FIG. 310. Resistances in Series

Conductors are said to be connected **in parallel** when a part only of the electricity passes through each conductor. This connection is represented by Fig. 311.

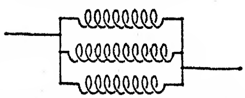


FIG. 311. Resistances in Parallel

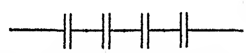


FIG. 312. Condensers in Series

Since the plates of a condenser are separated by a dielectric (insulator), electricity does not flow through a condenser. However, the terms "series" and "parallel" are applied to the connections of condensers also when they are arranged like the conductors described above. These connections are represented in Fig. 312 and Fig. 313.

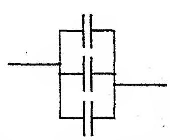


FIG. 313. Condensers in Parallel

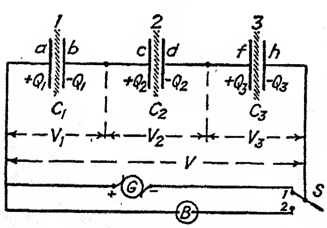


FIG. 314

366. Capacitance of a group of condensers in series. Given a group of condensers numbered 1, 2, 3, etc., connected in series (Fig. 314), it is required to find the capacitance of the group as a whole.

Consider first the charging process

(switch S on point 1). The generator G causes an electron to pass from plate a to plate h, thereby giving h a negative charge of 1 electron and leaving a with an equal + charge (1 proton). The same is true for each electron removed from a to h.

Hence the total + charge on a is equal to the total − charge on h; that is

$$+Q_1 = -Q_3 \quad \text{numerically.}$$

Since plates b and a are separated by a dielectric (insulator), plate b then becomes charged negatively and plate c charged positively by induction from a. Similarly, plates f and d become charged + and −, respectively, by induction from h. Therefore, by Sec. 334 (b) and (c),

$$\left.\begin{array}{ll} & +Q_1 = -Q_1 = +Q_2 \\ \text{and} & \\ & -Q_3 = +Q_3 = -Q_2 \end{array}\right\} \text{numerically.}$$

Hence the + and − charges on all members of a group of condensers in series are the same, or

$$Q_1 = Q_2 = Q_3. \tag{a}$$

Consider now the discharging process. Throw the high resistance switch S to contact point 2. This disconnects the generator G and connects the coulombmeter B, a ballistic galvanometer, in circuit with the condensers. The electrons of charge $-Q_3$ will flow through the coil of the coulombmeter, registering their quantity on the meter, and will neutralize the equal $+Q_1$ charge on plate a.

Being no longer bound, the electrons of plates b and d will flow back and neutralize the equal + charges on c and f, respectively; and the whole group will be completely discharged by the passage of $-Q_3$ through the coulombmeter B.

Hence the total charge Q of electricity on the whole group is, for all practical purposes, equal to $-Q_3$; for that is all that is measured during discharge.*

* Since, in charging or discharging a condenser, only the quantity of electricity on one plate changes sides, it is customary to speak of $+Q$ as "the quantity of electricity on a condenser." From the above it will be seen that it would be even better to say that $-Q$ is the charge on a condenser.

Therefore,

$$Q = Q_1 = Q_2 = Q_3. \tag{b}$$

Since the conductors are connected in series,

$$V = V_1 + V_2 + V_3 + \cdots. \quad \text{(c) by Sec. 350}$$

By the definition of capacitance, Eq. (271):

$$\left.\begin{aligned} C &\equiv \frac{Q}{V}; \quad V = \frac{Q}{C} \\ C_1 &\equiv \frac{Q_1}{V_1}; \quad V_1 = \frac{Q_1}{C_1} \\ C_2 &\equiv \frac{Q_2}{V_2}; \quad V_2 = \frac{Q_2}{C_2} \\ C_3 &\equiv \frac{Q_3}{V_3}; \quad V_3 = \frac{Q_3}{C_3} \end{aligned}\right\} \tag{d}$$

Substituting from Eq. (d) in Eq. (c):

$$\frac{Q}{C} = \frac{Q_1}{C_1} + \frac{Q_2}{C_2} + \frac{Q_3}{C_3} + \cdots$$

and since the Q's are all equal, we can divide through by Q, which gives the relation:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots. \tag{280}$$

367. Capacitance of a group of condensers in parallel. Given a group of condensers numbered 1, 2, 3, etc., connected in parallel (Fig. 315), it is required to find the capacitance of the group as a whole.

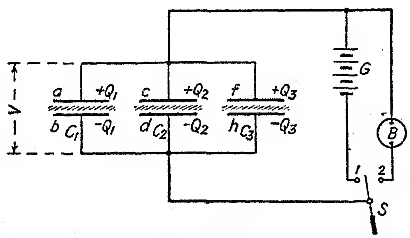


Fig. 315

Consider first the charging process. The generator G causes electrons to pass from the plates a, c, and f to the plates b, d, and h, charging the latter negatively and leaving the former with equal positive charges as shown.

If the switch S is then thrown so that the condensers may discharge through the coulombmeter B, the electrons composing the

charges $-Q_1$, $-Q_2$, and $-Q_3$ will all flow back through the meter, registering their amount on the meter and neutralizing the + charges on plates *a*, *c*, and *f*.

The total quantity Q on the group as a whole is therefore:

$$Q = Q_1 + Q_2 + Q_3 + \cdots, \tag{a}$$

for this is the amount that is measured during discharge.

As the condensers are connected in parallel, all the plates on either side form one conductor. Since a conductor having a static charge is an equipotential surface (Sec. 355), all the plates on one side come to the same potential, and all those on the other side come to a different potential. Consequently, when fully charged the potential difference between the plates of any one of the condensers equals the potential difference between those of any other condenser, and equals the potential difference V developed by the generator.

That is,

$$V = V_1 = V_2 = V_3 = \cdots. \tag{b}$$

By the definition of capacitance, Eq. (271):

$$\left.\begin{aligned} C &\equiv \frac{Q}{V}; \quad Q = CV \\ C_1 &\equiv \frac{Q_1}{V_1}; \quad Q_1 = C_1V_1 \\ C_2 &\equiv \frac{Q_2}{V_2}; \quad Q_2 = C_2V_2 \\ C_3 &\equiv \frac{Q_3}{V_3}; \quad Q_3 = C_3V_3 \end{aligned}\right\} \tag{c}$$

Substituting from Eq. (c) in Eq. (a), we have:

$$CV = C_1V_1 + C_2V_2 + C_3V_3 + \cdots.$$

But since, by Eq. (b) the V's are all equal, we may divide through by V, which gives the relation:

$$C = C_1 + C_2 + C_3 + \cdots. \tag{281}$$

368. Energy of a charged condenser. Let a condenser of capacitance C be charged by a generator capable of developing a potential difference V (Fig. 316).

When the charging begins, the potential difference between the plates of the condenser is 0; and when charging is complete, the potential difference between them is V.

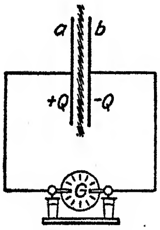


FIG. 316

Hence the **average** potential difference $\overline{V}$ against which the charge Q is "pumped" by the generator is: *

$$\overline{V} = \frac{0 + V}{2} = \frac{V}{2}.$$

Therefore, the work W done by the generator in transferring Q units of electricity from one plate to the other against the average potential difference of $\overline{V}$ is:

$$W = \overline{V}Q \qquad \text{by Eq. (264)}$$

$$= \frac{V}{2}Q$$

$$= \frac{1}{2}VQ \qquad (282)$$

$$= \frac{1}{2}CV^2 \qquad (283)$$

where W is in ergs if C is in statfarads and V in statvolts; and W is in joules if C is in farads and V in volts.

369. Electrostatic machines. Beginning with a frictional device invented by Otto von Guericke (1602–86) of Magdeburg and consisting of a sphere of sulphur rotating on an axis and rubbed by the dry palm of the hand, various machines were devised for developing large charges of static electricity. Machines employing the method of induction, commonly called **influence machines,** proved to be capable of developing larger charges and were less affected by atmospheric conditions than machines depending upon friction alone.

The **electrophorus** is the simplest of the induction, or influence, machines. It consists usually of a pie pan filled with sealing wax,

* This is true only when the variables (in this case, Q and V) are connected by a linear relation. Here $Q = CV$, which is the equation of a straight line. (Cf. case of $\bar{v}$, Sec. 27.)

or a plate of hard rubber, and a somewhat smaller metal lid having an insulating handle.

The cake of sealing wax is electrified negatively by being rubbed with fur (Fig. 317a). The lid is then lifted by the handle and placed upon the wax. Being warped (which is necessary), it touches the sealing wax at only a few points; and since the wax is a nonconductor, the amount of negative electricity which transfers from the wax to the lid by conduction is negligible. By induction, electrons are repelled to the upper surface of the lid, leaving

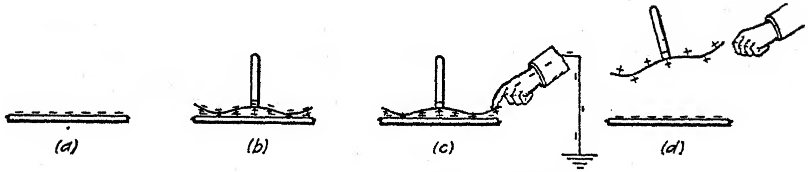


FIG. 317. Charging an Electrophorus

an excess of positive charges on its lower surface (Fig. 317b). The lid is now touched with the finger, or otherwise grounded. The electrons are then repelled to the earth, and a bound positive charge remains on the lid (Fig. 317c). If the lid is now lifted from the sealing wax by its insulating handle, the fact that it has a considerable charge is demonstrated by a vigorous spark when the knuckle is brought near (Fig. 317d). Without further rubbing with the fur and without apparent diminution of the original charge on the sealing wax, a great many sparks may be obtained by repeating the last three steps of the process. Loss of charge, which is negligible, occurs only at the points of contact: **in fact, contact is not necessary.**

The sparks evidence the dissipation of energy, since with them are associated heat, light, and sound. The question naturally arises: What is the source of so great a supply of energy? It cannot be the original rubbing with the cat fur, because the sealing wax does not lose its charge. The answer is found in the fact that the energy of each spark is stored in the system **as the lid is lifted,** and represents the work done in pulling the protons on the lid away from the attractive electrons on the sealing wax. Hence, as we should expect, the spark is more energetic when the lid is lifted a foot than when it is lifted only a few inches.

370. The Wimshurst machine has proved to be one of the best electrostatic generators of the influence type. Figure 319 is a diagrammatic drawing of the machine shown in perspective in Fig. 318, both figures being lettered alike.

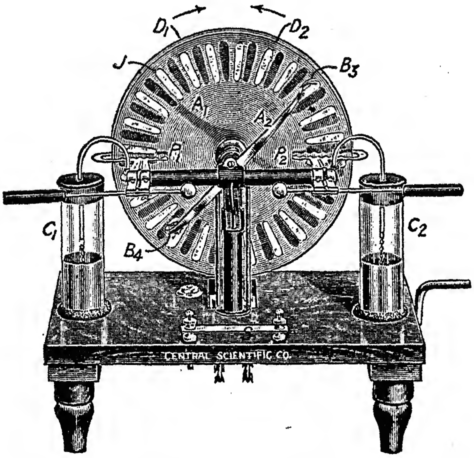


FIG. 318. The Wimshurst Machine. (Courtesy of Central Scientific Company)

The machine consists of two circular plates, D_1 and D_2, of glass or other insulating material, rotating in opposite directions on the same shaft. In Fig. 319 these plates are represented by the stippled circles. They carry on their outboard faces a large number of tin-foil sectors t_1, t_2, etc., which act as "inductors" and "carriers" of electricity.

The action of the Wimshurst machine is essentially that of a continuously acting electrophorus, and is as follows.

Because of the ions always present in the atmosphere, some sector J will have an excess of electrons. When this sector passes opposite the brush B_1 it repels electrons to the end B_2 of the equalizing arm A_1, leaving the sector on plate D_1 and opposite to J charged positively with a bound + charge. As the plates revolve, the negative charges at J and B_2 are carried to the left, and the positive charge at B_1 is carried to the right. When it comes opposite the

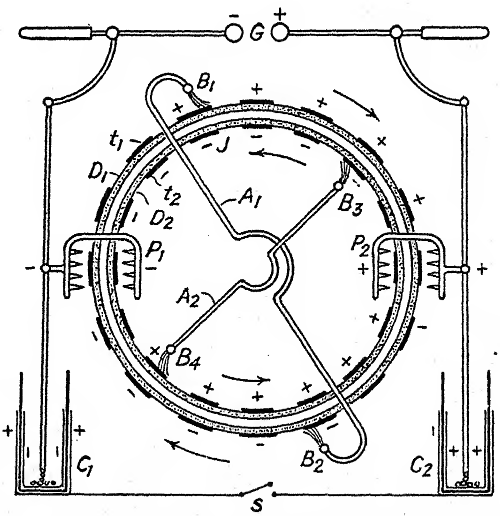


FIG. 319. Diagram of Wimshurst Machine

brush B_3, it induces a − charge on the sector of D_2 that is in contact with B_3, leaving a bound + charge on the sector of D_2 that is in contact with B_4.

When these sectors of D_2 pass out of contact with B_3 and B_4, they carry their − and + charges to the left and right, respectively. On arriving at J this − induces a + charge on the sector under B_1, as before; and the whole process is repeated indefinitely.

It will be seen that both plates are carrying + charges to the right and − charges to the left. These charges are drawn off from the sectors by the sharp points P_2 and P_1, and accumulate on the inner coatings of the condensers C_2 and C_1, respectively.

If the switch S is closed, C_1 and C_2 are in series; when it is open, they are virtually in series with a third condenser consisting of the outer coatings of C_1 and C_2, separated by the wooden base of the machine as a (poor) dielectric. In the former case the capacitance of the group of condensers is greater than in the latter case; and a greater quantity of electricity is required to bring the two sides of the system to a potential difference sufficient to cause the air in the gap G to break down. Consequently, when the switch is closed the sparks will be larger and less frequent than when it is open, for it takes longer to build up the greater charge necessary to raise the potential difference to the break-down value.

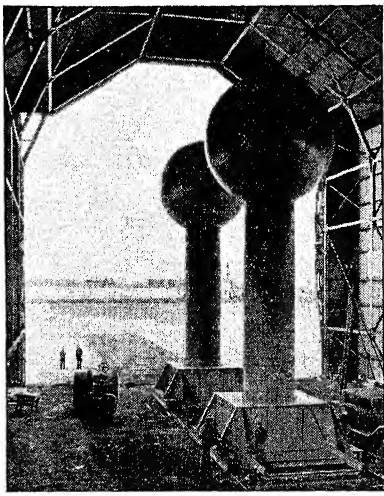


FIG. 320. Van de Graaff Electrostatic Generator. (Courtesy of Professor R. J. Van de Graaff)

371. The Van de Graaff electrostatic generator. The largest electrostatic generator that has been built thus far is one constructed in 1933 at Round Hill, Mass., by L. C. and C. M. Van Atta, D. L. Northrup, and R. J. Van de Graaff for the Massachusetts Institute of Technology.* Its great size may be seen by comparison with the men and the automobile standing beside it,

* *Physical Review* (May 15, 1936).

in Fig. 320; and its operation may be understood from the diagrammatic sketch of Fig. 321.

The positive terminal P and the negative terminal N are spheres of aluminum 1/4 in. thick and 15 ft in diameter. These are supported on cylinders of insulating Textolite 5/8 in. thick, 6 ft in diameter, and 22 ft high. Each of these towers is mounted

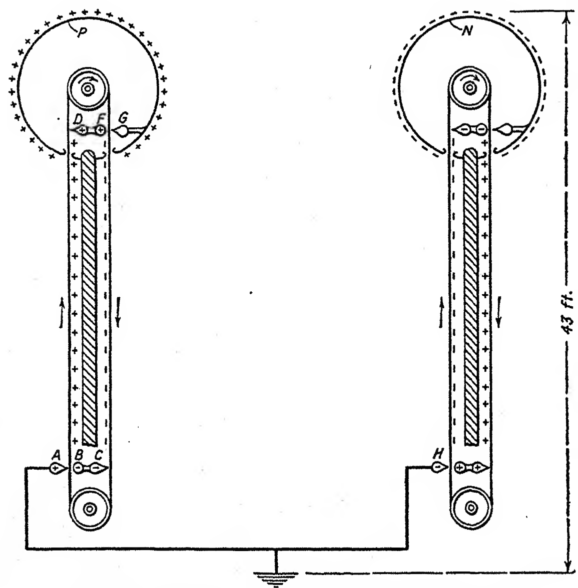


FIG. 321. Diagram of Van de Graaff Generator.

on a steel truck which can be moved along a 14-ft gauge railway track. The over-all height is about 43 ft. In each vertical cylinder run two paper belts 47 in. wide which serve to separate the charges, just as the glass disks and the sectors do in a Wimshurst machine. For clearness, only one belt is shown in each diagram.

In actual operation, charges are "sprayed" on to the belts from a 20,000-volt kenotron rectifying set; but the machine may be operated as a "self-excited" unit. The latter method, being less complicated, will be described here. The diagram is further simplified by considering the charges to be carried on the inner surfaces of the belts, whereas in the actual machine they are carried on the outer surfaces.

Friction of the dry paper belts on the motor-driven pulleys will develop some charge. Suppose a + charge carried upward by the

belt on the left, as shown. As it passes point D it will be neutralized by electrons attracted from the point of D, leaving F with a + charge. This will induce a − charge on the inside of the belt and a + charge on its outside. The latter will attract electrons from the point of G, leaving G and the sphere P positive. The corresponding − charge, bound on the belt by the + charge on F, will be carried downward. On reaching C, this − charge repels an equal − charge to B and neutralizes the + on the point of C. The − charge on B induces another + charge on the belt, which is carried upward as before, and repels electrons to the point of A, whence they pass to H and "spray" upon the upward moving belt. The belt in the tower on the right carries − charges to the terminal N and brings down + charges.

The process of separating + and − charges is continuous in both towers, and the charges accumulate on the outer surfaces of the spheres with consequent increase in the potential difference between them. Within the spheres the field intensity is zero, and men may work in them while the machine is in operation.

Built to be used as an "atom smasher" in nuclear research, the generator develops a potential difference of 5,100,000 volts, at which it can deliver a current of 1.1 milliamperes.

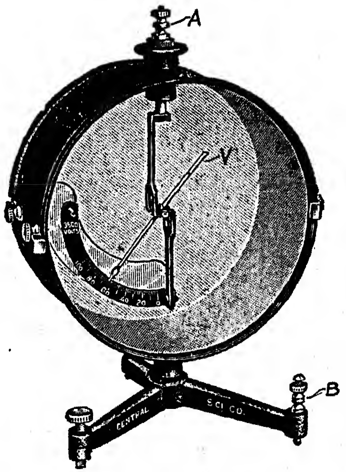


FIG. 322. Electrostatic Voltmeter. (Courtesy Central Scientific Co.)

372. Electrostatic voltmeters. When the metallic case of a gold leaf electroscope is connected to the earth, and the knob to an electrified body, the divergence of the leaves is a measure of the potential difference between the body and the earth. If a scale, which may be calibrated in volts, is provided to measure the divergence of the leaves, a gold leaf electroscope becomes an electrostatic voltmeter.

The *Braun electrostatic voltmeter* (Fig. 322) is of this type. The gold leaf is replaced by a light aluminum vane, delicately balanced, which deflects from its vertical supporting rod when the

two are charged from a body connected to the binding post *A*. The metallic case is connected to the earth or to the other body at *B*.

The electrostatic voltmeter has the advantage that **it requires no current**; and the same scale is correct for systems in which the current is either direct, or alternating of low frequency.

373. Millikan's oil-drop experiment.* One of the most important of the natural constants is the quantity of electricity represented by one electron. The magnitude of this charge (commonly denoted by *e*), in fact the independent existence of such charges, was determined by Professor R. A. Millikan at the University of Chicago in a series of most searching and refined experiments extending over the period from 1909 to 1917.

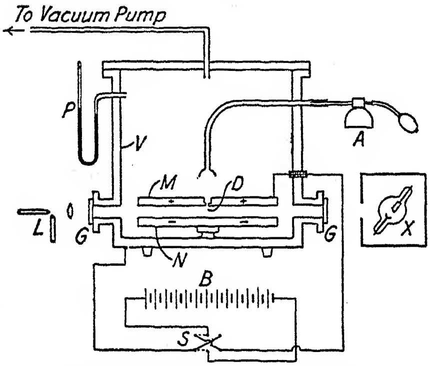


FIG. 323. Millikan's Oil Drop Experiment

His procedure, known as the "balanced-drop method," was substantially as follows.

An air condenser, consisting of two circular plates *M* and *N* (Fig. 323) separated by a layer of air about 1.5 cm thick, is mounted in a vessel *V* within which the pressure *P* may be regulated by a vacuum pump. *M* and *N* are connected through a reversing switch *S* to a storage battery *B* of 5000 to 10,000 volts.

Droplets of oil, having diameters of the order of 1/1000 mm, are sprayed into *V* by means of the atomizer *A*, and a droplet *D* will occasionally find its way through a pinhole at the center of *M* into the space between *M* and *N*. Such droplets are too small to be seen by the unaided eye, and are viewed by means of a short-focus telescope (not shown) which is set to look perpendicularly to the paper at *D*.

In general, these droplets are strongly charged by friction in the spraying process. If the plates *M* and *N* are uncharged, the

* R. A. Millikan, *Electrons (+ and −), Protons, etc.* (University of Chicago Press), p. 57.

drop will fall freely under the action of gravity; and from its speed of fall, its weight mg can be computed by Stokes' law.

Suppose a drop under investigation is negatively charged and has nearly reached the plate N. If the switch S is then thrown so as to make the upper plate + and the lower plate −, the negatively charged droplet will be drawn upward. The potential difference between the plates could then be adjusted until the weight was exactly "balanced" by the upward force Fq, where F is the uniform field intensity between the plates and q is the charge on the droplet. We should then have

$$Fq = mg$$

from which q could be found since the other factors are known.

It is easier, however, to observe the speed with which the droplet rises and falls as the charges on the plates are reversed in sign, and to calculate q from these and other easily obtained data. By reversing the charges on M and N just before the droplet reached either plate, Millikan was able to keep **the same droplet under observation for several hours.**

The speed depends upon the charge that the droplet carries, and a sudden change of speed means that the droplet has captured a + or a − ion from the air. By means of an x-ray tube X, the air between the plates may be ionized and the charge on a droplet adjusted almost at will, from a certain smallest value to 150 times this value.

Thousands of observations were made by many observers, and invariably the charge, whether + or −, acquired by the droplet was an integral multiple of the smallest charge ever (but frequently) observed. There were no fractions such as halves, thirds, sevenths, etc. Hence we must conclude that **electricity is granular, or atomic, in structure,** the unit being this smallest quantity, which is called the charge of one electron e. The value of this **natural unit charge** as obtained by Professor Millikan was

$$\begin{aligned} e &= 4.770 \pm 0.005 \times 10^{-10} \text{ statcoulomb} \qquad (284) \\ &= 1.590 \times 10^{-19} \text{ coulomb.} \end{aligned}$$

The experiment was made also by spraying glycerine and mercury instead of oil, and the same value of e was always obtained. The experiment therefore demonstrated that

1. **There is a definite minimum quantity, or natural unit, of electricity** that occurs in nature.

2. All charges are made up of a whole number of these unit charges.

3. These atomic packages of electricity are the same quantity whether obtained by friction, as cathode rays, as beta-rays of radium, by ionization of a gas by x-rays, or from any known source.

4. The electron is the same whether obtained from a conductor (mercury), a nonconductor (oil), or from a semi-conductor (glycerine).

374. Piezoelectricity. Certain crystals, notably Rochelle salt, quartz, tourmaline, and boracite, have the property, when com-

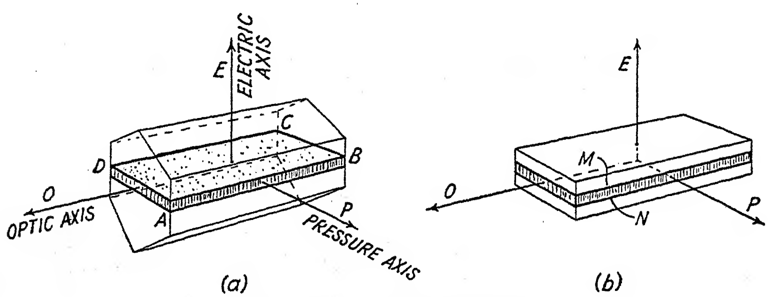


FIG. 324. X-Cut Quartz Plate

pressed along a certain axis, of developing opposite charges at definite parts of the crystal. If subjected to tension instead of compression, the signs of the charges are reversed. Such charges are called *piezoelectricity* (pressure electricity). They were discovered by J. and P. Curie in 1880.

As an example, let a plate $ABCD$ be cut from a natural quartz crystal (which is hexagonal), as shown in Fig. 324a, the optic axis being lengthwise of the crystal; and let the surfaces of the plate be ground approximately plane and parallel. If this plate is then placed between two metal electrodes M and N (Fig. 324b) and pressure applied along the axis P, charges of opposite sign will appear on the faces in contact with the plates M and N.

In the reverse effect, if the plates M and N are given charges

of the proper sign, the crystal plate will expand along the axis E and contract along the axis P; but there will be no change along the optic axis O. Reversal of the charges produces contraction along axis E and expansion along axis P. This phenomenon is known as **electrostriction.**

If charges alternating at the **natural frequency** of mechanical vibration of the crystal are impressed upon the electrodes M and N, the crystal may be kept in sustained vibration. Such vibrating crystals are our best source of constant frequency, Fig. 325.

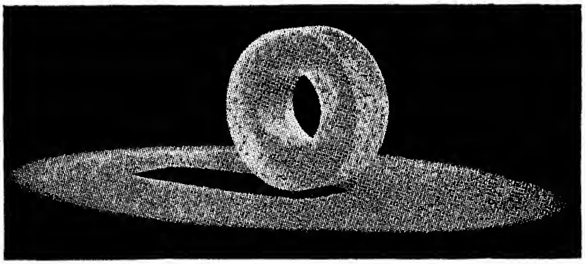


FIG. 325. Quartz Crystal with Zero Temperature Coefficient Used in Crystal Clock. (Courtesy of Dr. W. A. Marrison)

W. G. Cady has applied them to the control of the frequency of broadcasting; and Dr. W. A. Marrison of the Bell Telephone Laboratories has used them instead of a pendulum in a clock of the highest precision.

Pyroelectricity. Crystals that show piezoelectrification have also the property, if heated above any temperature at which they are electrically neutral, of developing equal and opposite charges on two opposite faces. If cooled below the temperature of neutrality, the charges again appear with reversed signs. The phenomenon is called **pyroelectrification** and has been known since 1700.

PROBLEMS

1. Charges of $+160$ and -100 esu, respectively, are 16 cm apart in a medium whose dielectric constant is 2.4. What force acts upon each?

2. Charges of $+200$ and -150 esu, respectively, are 20 cm apart in a medium whose dielectric constant is 2.5. What force acts upon each?

3. Two small spheres each weighing 0.1 gm and having equal charges are suspended from the same point by silk fibers 90 cm long. If the spheres are kept 6 cm apart by repulsion, what is the charge on each?

4. Two small spheres, each weighing 0.15 gm and having equal charges, are

suspended from the same point by silk fibers 20 cm long. If the threads make an angle of 10° with each other, what is the charge on each?

5. Two small conducting balls suspended by insulating threads have the same diameter and mass. They are charged with −20 and +60 esu, respectively, and are then allowed to touch, after which they stand apart a distance of 4 cm. What force does each exert on the other?

6. What is the difference of potential between *A* and *B* if it requires 1200 ergs of work to carry a charge of +4 esu from *B* to *A*?

7. What work is done in carrying a charge of +5 esu from a place where the potential is −10 esu to where it is +50 esu?

8. Charges of +96 and −48 esu are 8 cm apart. What is the field intensity and the potential at the point midway between them?

9. In air, charges of 40 and 50 esu are placed at the vertices *A* and *B*, respectively, of a triangle *ABC* whose sides *AB*, *BC*, and *AC* are 10, 8, and 6 cm, respectively. Find the field intensity at *C*.

10. Charges of +64 and −48 stc are at diagonal corners of a square each edge of which is 4 cm long. Compute the field intensity at one of the other corners.

11. A rectangle *ABCD* has sides *AB* and *BC* 12 cm and 8 cm long, respectively. Charges of +200, −100, +36, and −72 esu are placed at points *A*, *B*, *C*, and *D*, respectively. Find the potential at the middle point *M* of *CD*.

12. Two point charges of +36 and −48 esu, respectively, are 5 cm apart in air. What is the field intensity and the potential at a point 3 cm from the former point and 4 cm from the latter?

13. A condenser has a capacitance of 2 microfarads. With how much electricity must it be charged in order to make the potential difference between its plates 110 volts?

14. A parallel plate condenser is immersed in a medium whose dielectric constant is 2.4. What is the capacitance of the condenser if it has 17 plates of conductor, each 8 cm × 10 cm and 0.4 cm apart?

15. A parallel plate condenser is immersed in oil whose dielectric constant is 2.5. What is the capacitance of the condenser if it has 21 plates of area 30 cm^2 each and the plates are 0.1 cm apart?

16. A parallel plate condenser is immersed in a medium whose dielectric constant is 1.8. What is the capacitance of the condenser if it has 19 plates of conductor, each 10 cm × 15 cm and 0.2 cm apart?

17. A Leyden jar of 15 esu capacitance has the potential difference between its coatings raised from −5 esu to +25 esu. Calculate the total work required.

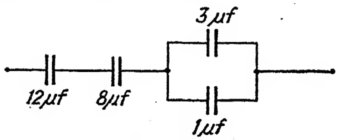
18. How much energy is possessed by a condenser which has a capacitance of 8 statfarads when the potential difference is 50 stv?

19. Compute the combined capacitance of three condensers having capacitances of 20, 40, and 50 statfarads, respectively, when connected (1) in series, and (2) in parallel.

20. Compute the combined capacitance of three condensers having capacitances of 15, 20, and 30 statfarads, respectively, when connected (1) in series; (2) in parallel.

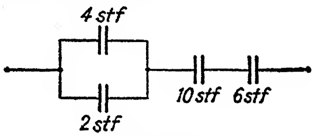
21. Compute the combined capacitance of three condensers having capacitances of 10, 20, and 30 statfarads, respectively, when connected (1) in series, and (2) in parallel.

22. Find the capacitance of the group:



PROB. 22

23. Find the capacitance of the group:



PROB. 23

CHAPTER XXV

MAGNETISM

375. Historical. Thales of Miletus, one of the "seven wise men" of early Greece, who, according to tradition, knew that amber could be electrified by rubbing, is also credited with knowing (600 B.C.) that a certain kind of iron ore, now called magnetite (Fe_3O_4), would attract iron.

Pieces of the mineral exhibiting this property were called magnets, or **natural magnets,** because they were found in Magnesia in Asia Minor. They are also called lodestones, i.e., "leading stones."

The first scientific work on magnets was *De Magnete*, published in 1600 by William Gilbert, physician-in-ordinary to Queen Elizabeth. He was an ingenious experimenter, and established most of the commonly known facts about magnets except those associated with electric currents.

376. Artificial magnets may be made by properly stroking a piece of iron or steel with a natural magnet, or by placing the piece to be magnetized in a coil of wire in which electricity is flowing. Much stronger magnets are produced by the latter than by the former method.

Magnets made of soft iron are *temporary*, retaining their magnetism for a short time only; but if made from hard, cobalt, or tungsten steel, they are practically *permanent.*

Bar magnets (Fig. 326a) are preferred for many uses; but by bending the bar before magnetization into a U-shaped **horseshoe**

FIG. 326. Permanent Magnets

magnet (Fig. 326b), to decrease the distance between its ends, much greater effect is secured, as for lifting purposes.

377. Magnetic poles. If a magnet is lifted out of a pile of filings or small tacks, these bits of iron will be found to cling in clusters about two regions, usually near the ends (Fig. 327).

Gilbert called these "points" at which the magnetism appeared to be concentrated the **poles** of the magnet. But they are regions

FIG. 327. Iron Filings Cluster about Magnet Poles. (Chicago Apparatus Co.)

rather than points, for they cannot be located with precision, and their positions shift somewhat as two magnets approach each other.

Again, if a number of magnets are suspended so as to turn freely about a vertical axis, and are at a considerable distance from one another, it will be found when they come to rest that all of them point approximately north and south. The pole toward the north is called the north-seeking pole, or **north pole,** of the magnet; and the other one, the south-seeking pole, or **south pole.** North poles are called also **positive** poles; south poles, **negative.**

A magnet always has at least two poles; a single isolated magnetic pole cannot be obtained. Even if a magnet is broken into

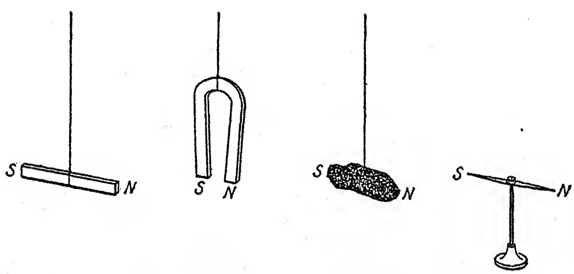


FIG. 328. Suspended Magnets

very small fragments, each fragment still exhibits two poles. This constitutes an **important difference** between magnetic poles and electric charges: a positive pole cannot be separated from an associated equal negative pole, whereas positive and negative charges of electricity can be separated readily.

That **the two poles are of equal strength** may be shown as follows. Let a magnet be placed on a piece of cork and that in turn

on the surface of a dish of water. The magnet will take up a north and south direction by a motion of pure rotation, i.e., without any longitudinal motion. If either pole were the stronger, the magnet would drift in the corresponding direction. Since it does not drift, the poles must be of equal strength.

378. Du Fay's law of magnetism. After the magnets of Fig. 328 have come to rest, let their north and south poles be marked. If any one of the magnets is then removed from its support and its poles are presented successively to the poles of those that remain, it will be found that any north will repel any other north, and that any south will repel any other south. Furthermore, any north will attract any south, and vice versa. From such data, Du Fay gave in 1745 the law which bears his name: **Like poles repel each other; unlike poles attract each other.**

379. Coulomb's law of magnetism. By means of a torsion balance, as in Fig. 329, Coulomb obtained in 1785 the quantitative relation for the force exerted by two magnetic poles upon each other in air. He used long slender magnets so that the effect of the more remote poles could be neglected. Subsequently it was found that the medium in which the poles are immersed has an effect upon the force. The law is as follows.

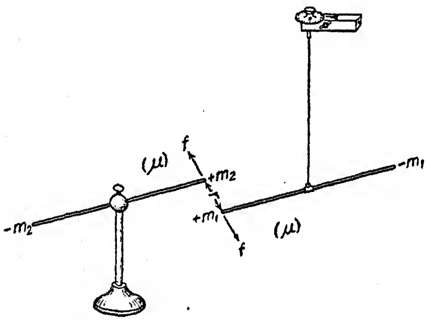


FIG. 329. Illustrating Coulomb's Law

The force f of attraction or repulsion between two isolated point magnetic poles varies directly as their pole strengths m_1 and m_2 and inversely as a property μ of the medium and as the square of the distance r between them.

In symbols,

$$f \propto \frac{m_1 m_2}{\mu r^2}$$

or

$$f = C'' \frac{m_1 m_2}{\mu r^2} \qquad \text{for any units} \qquad (285)$$

where C'' is a constant depending upon the choice of units.

The forces act along the line joining the two poles.

The law is an *ideal* one, since it is not possible to secure isolated point poles, and when applied to actual magnets it gives results that are only approximately correct. But it is most useful as a basis for the definitions and theory that follow. It will be noted that Coulomb's law for magnetic poles has exactly the same form as his law for electrostatic charges, Eq. (255).

The effect of those media in which it is possible to suspend magnets is so slight that it is not feasible to determine the permeability μ of a medium by this relation. Permeability is therefore defined in Sec. 454, where the method of determining it is described.

380. Electromagnetic units. Thus far no units of magnetic pole strength have been adopted, the units used having been entirely arbitrary. Since C'' depends upon the choice of units, we now proceed to choose units so as to eliminate C'' by making it unity.

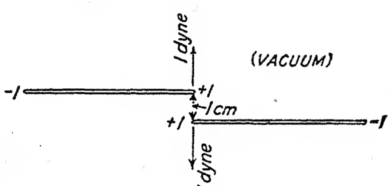


FIG. 330. Unit Magnet Poles

1. The **unit of permeability** μ is defined as the permeability of a vacuum.

2. The **unit magnetic pole** is defined as a pole which will repel an equal like pole at a distance of one centimeter in a vacuum with a force of one dyne.

We may imagine an experiment (Fig. 330) in which the quantities are those chosen in these two definitions. Substituting these values in Coulomb's law, Eq. (285), we have:

$$1 \text{ dyne} = C'' \frac{(1 \text{ unit pole})(1 \text{ unit pole})}{(1 \text{ for vacuum})(1 \text{ cm})^2}$$

whence

$$C'' = 1 \qquad \text{for these units.}$$

The above units are the first of the **absolute electromagnetic system** (emu), from which later on we shall obtain the practical system of units used in the majority of everyday electrical measurements.

Coulomb's law may now be written in the simplified form:

$$f = \frac{m_1 m_2}{\mu r^2} \qquad \text{provided emu are used.} \qquad (286)$$

For air, in which most magnetic experiments are made, $\mu = 1.0000004$; and it is approximately unity for most materials. Hence, in the majority of cases, μ may be omitted from the formula.

381. Magnetic field. A magnetic field is a region in which magnetic effects may be detected. We do not always know where the magnetic field originates. For example, in Fig. 328, the fact that the magnets all orient themselves in a direction approximately north and south indicates that they are in a magnetic field, which we call the earth's field—but we do not know definitely what causes this field. Every magnet produces about it a magnetic field which theoretically extends to infinity, but its practical extent is seldom beyond a few meters.

When iron filings are scattered over a sheet of paper under which a magnet lies, and the paper is tapped to relieve friction,

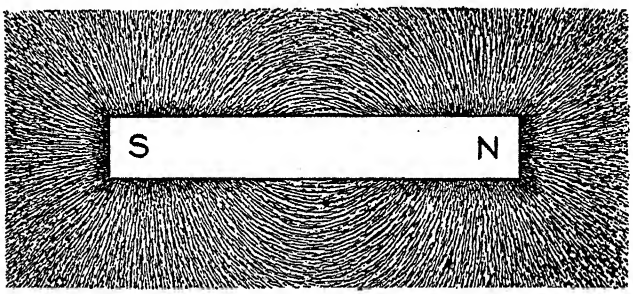


FIG. 331. Field about a Magnet

the filings arrange themselves in symmetrical curves about the magnet, being most concentrated where the field is strongest (Fig. 331). These lines present to the eye an excellent picture of the distribution of the magnetic field, and gave Faraday the idea of "lines of force," which he used most effectively in explaining many magnetic phenomena.

Each tiny bit of iron, when it falls into the magnetic field, becomes a magnet by induction and moves into a position of equilibrium under the forces present. Exactly similar curves may be traced out by using a small compass, a dot being placed at each end of the compass needle and the needle then moved up until its south end comes to the point previously placed opposite its north end.

A **line of magnetic force** is therefore defined as the path followed by a north magnetic pole when free to move in a magnetic field. The **direction of the field** at a point is taken as the direction of motion of the north pole as it describes the line of force through that point.

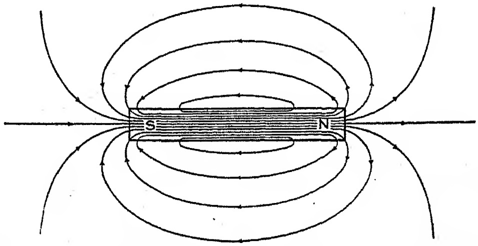


FIG. 332. Magnetic Lines of Force Are Continuous

Lines of force never cross one another. They appear to come out of the north end and re-enter the south end of a magnet. They are believed to be closed curves continuing through the magnet from its south to its north end (Fig. 332), because they do so in the case of a solenoid (see Fig. 345); but this cannot be verified in the case of a solid magnet.

In the following sections, this qualitative method of depicting a magnetic field is made quantitative.

382. Test pole. The test pole by means of which a magnetic field is investigated is taken to be a **unit positive pole.** As it would be difficult to prepare a pole of exactly unit strength, any pole of known strength may be used; but results are reduced to what they would have been if a unit positive pole had been used.

383. Magnetic field intensity. The intensity, or strength, H of a magnetic field at a point is defined as **the force in dynes that acts upon a unit + pole when placed at the point in question.** Hence, if f is the force in dynes that acts upon a pole of strength m_1 unit poles when placed at a certain point, then the field intensity at that point is:

$$H \equiv \frac{f}{m_1} \frac{\text{dynes}}{\text{unit pole}}, \text{ or oersteds.} \qquad (287)$$

If we know the strength m of the point pole that produces the field under consideration, we can now find an expression for the intensity of that field. For when a pole of strength m_1 is placed at a distance r cm from the pole of strength m which produces the

field, it will be acted upon by a force f as given by Coulomb's law:

$$f = \frac{mm_1}{\mu r^2}$$

which, substituted for f in Eq. (287), gives:

$$H = \frac{m}{\mu r^2} \frac{\text{dynes}}{\text{unit pole}}, \text{ or oersteds.} \tag{288}$$

The direction of H, which is also called **the direction of the field,** is the direction of the force that acts upon the $+1$ test pole. Hence H is a **vector** quantity.

Magnetic field intensity may be represented by drawing *lines of force*, just as was done for electric field intensity (Sec. 344).

The **unit of** H is obtained from its defining equation (287), where it is seen that H is unity when f is 1 dyne and m_1 is 1 unit pole. This unit of field intensity is called the **oersted** after Hans Christian Oersted, professor of physics at the University of Copenhagen, who discovered in 1819 that a magnetic field surrounds a wire in which electricity is flowing. We have therefore:

$$1 \text{ oersted} \equiv \frac{1 \text{ dyne}}{1 \text{ unit pole}}.$$

That is, **an oersted is the intensity of a magnetic field at a point when a unit magnetic pole placed at the point in question is acted upon by a force of one dyne.**

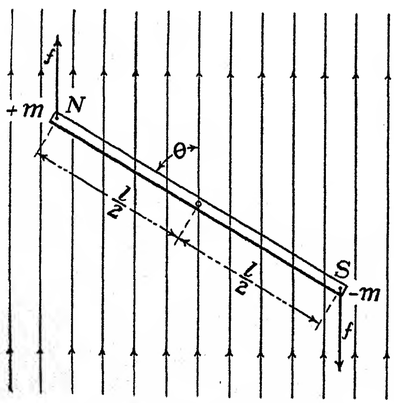


FIG. 333. Magnetic Moment

384. Magnetic moment. Since it is impossible to secure a magnetic north pole unaccompanied by an equal south pole, and since magnetic poles cannot be located as definite points, experiments with actual magnets must be carried out with the two poles attached together. Such a combination is called a **magnetic doublet,** or **dipole,** and has a definite magnetic moment. This is a useful property and is readily

determinable without knowing the pole strength and distance between the poles separately.

Consider an ideal magnet *NS* (Fig. 333) of length l cm and having poles of strength m at its ends. Let this magnet be suspended at its center O in a uniform magnetic field of intensity H by a torsion fiber perpendicular to the paper. Then each end of the magnet is acted upon by a force f as shown, where

$$f = Hm \qquad \text{by Eq. (287),}$$

and these forces produce a couple whose torque is:

$$\begin{aligned} \mathfrak{T} &= fl \sin \theta \\ &= Hml \sin \theta. \end{aligned}$$

The product ml is called the *magnetic moment* M of the magnet:

$$M \equiv ml \tag{289}$$

so that

$$\mathfrak{T} = HM \sin \theta.$$

Here it will be seen that when $H = 1$ oersted and $\theta = 90°$, $M = \mathfrak{T}$. Hence, **magnetic moment is measured by the torque necessary to hold a suspended magnet at right angles to a magnetic field whose intensity is one oersted.** It will be noted that M may be found without knowing m and l separately.

385. Magnetic flux density: total flux. If a piece of magnetizable material such as iron is placed in a magnetic field of intensity H oersteds, more than H lines/cm² are required to carry out the convention of representing the state of magnetization within the material by lines, on account of the magnetic poles induced in the material itself. The analysis is quite similar to that of Sec. 346 in electrostatics.

The number of lines per square centimeter required is called the magnetic induction, or **flux density** B. This is found as described in Sec. 454, where we obtain the relations:

$$\mu \equiv \frac{B}{H} \quad \text{and} \quad B = \mu H. \tag{290}$$

The **unit of flux density** is logically **one line per square centimeter.** This unit has been officially named the **gauss** after Carl

Frederick Gauss, who, with William Weber, originated the absolute system of units in 1832.

$$1 \text{ gauss} \equiv 1 \text{ line/cm}^2.$$

If the flux density B is uniform over a certain area A perpendicular to the flux density, the total number of lines through the area is:

$$\Phi \equiv BA \tag{291}$$

and Φ is called the **total flux** through the area.

The **line** is the logical unit of flux in the electromagnetic system. It is officially named the **maxwell** after James Clerk Maxwell, who formulated the electromagnetic theory of light in 1864.

$$1 \text{ maxwell} \equiv 1 \text{ line of magnetic flux.}$$

The **practical unit of magnetic flux** is called the **weber** after William Weber, mentioned above.

$$1 \text{ weber} \equiv 10^8 \text{ maxwells, or lines.}$$

The word "flux," which implies a flow of something, was introduced into the literature of magnetics in the early days when magnetism was thought of as a fluid. It was an unfortunate choice, for we are now sure that magnetism does not flow like electricity. Magnetic phenomena are static, and the very exact analogy between many electrostatic and magnetic effects should be carefully noted.

By reasoning similar to that of Sec. 347 in electrostatics, we may show that at a distance r cm from a point magnetic pole of strength m the **flux density due to that pole** is:

$$B = \frac{m}{r^2} \text{ gausses, or lines/cm}^2. \tag{292}$$

Likewise, following Sec. 349, **Gauss' law** for magnetic flux becomes: The total flux from a pole of strength m unit poles is

$$\Phi = 4\pi m \text{ maxwells, or lines.} \tag{293}$$

386. Paramagnetic and diamagnetic substances. Prior to 1845, it was thought that only iron and nickel and their alloys were susceptible to magnetization. To Faraday this "appeared too

extraordinary to be probable." Accordingly, he built very powerful electromagnets (Sec. 455) and found, as he suspected, that apparently all substances are affected by magnetization. He divided them into two classes:

1. *Paramagnetic substances*, which have a permeability greater than that of a vacuum; i.e., $\mu > 1$.

2. *Diamagnetic substances*, which have a permeability less than that of a vacuum; i.e., $\mu < 1$.

These names, suggested by Whewell, arose from the fact that in a strong magnetic field paramagnetic substances in the form of rods tend to turn parallel to the direction of the field, whereas rods of diamagnetic substances tend to turn across (cf. diameter) the field.*

Paramagnetic substances are attracted by a magnet; i.e., they tend to move into the strongest part of the field. The reverse is true of diamagnetic substances.

A significant fact is that the most markedly paramagnetic elements—iron, cobalt, and nickel—**occur consecutively in the periodic table,** their atomic numbers being 26, 27, and 28, respectively. These elements have permeabilities that are very much greater than those of other paramagnetic substances, and which obey different laws. They are, therefore, placed in a special class of **ferromagnetic bodies,** which comprises substances whose permeabilities approximate that of iron (ferrum) and behave similarly. To this class belong certain alloys, notably Heusler's alloy (15 Al, 61 Cu, 24 Mn), Permalloy (78 Ni, 22 Fe), Perminvar (45 Ni, 25 Co, 30 Fe), and Alnico (12 Al, 20 Ni, 5 Co, 63 Fe).

The great majority of other substances have permeabilities so nearly unity that for most purposes they behave like a vacuum or air with respect to a magnetic field, and are commonly referred to as being **nonmagnetic.** If, however, observations are carried to six or eight significant figures, manganese, aluminum, air, and oxygen are found to be paramagnetic; whereas copper, tin, zinc, wood, paper, glass, rubber, body tissue, water, nitrogen, hydrogen, and helium are diamagnetic. *Bismuth is the most diamagnetic substance known.*

* Theory predicts that in a perfectly uniform magnetic field, which is extremely difficult to secure, the latter statement does not hold true.—A. G. Webster, *Theory of Electricity and Magnetism* (New York, The Macmillan Co., 1897), p. 371.

Substance	Permeability	Substance	Permeability
Soft wrought iron	3750	Vacuum	1. (by definition)
Cast semi-steel	1400	Air	1.0000004
Machine steel	460	Platinum	1.00002
Nickel	296	Water	0.999991
Cobalt	177	Bismuth	0.9998
Heusler's alloy	115		

The **influence of the medium** on the magnetic behavior of a substance was first observed by Faraday. A body behaves as paramagnetic or diamagnetic according as its permeability is greater or less than that of the surrounding medium. Thus a tube of ferric chloride places itself parallel to the magnetic field when suspended in a ferric chloride solution of less concentration, but it turns crosswise to the field in a solution of concentration greater than that in the tube.

387. Terrestrial magnetism. The fact that magnets point nearly north when properly suspended was known in Europe as

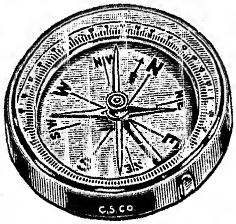


Fig. 334. Compass. (Courtesy Central Scientific Co.)

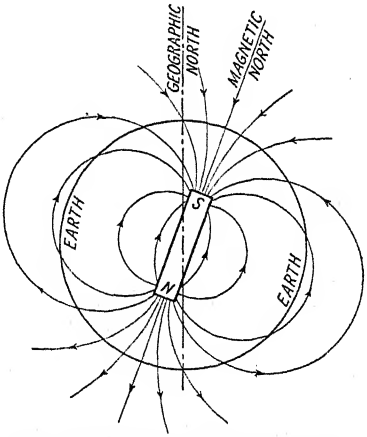


Fig. 335. Earth's Magnetic Field

early as the twelfth century, and probably earlier in China. A magnet mounted so as to rotate freely in a horizontal plane, for use in determining direction, is called a **magnetic compass** (Fig. 334). Columbus used such a compass on his first voyage to America.

This north-pointing property of magnets is easily accounted for by considering the earth itself to be a great magnet. It is usual to call the pole near the geographical north pole, the "earth's north magnetic pole"; although obviously it must be magneti-

cally a south pole, since it attracts what we have defined as the north pole of a magnet.

The earth's north magnetic pole was located in 1935 at latitude 71° N, longitude 96° W; and its south magnetic pole, at latitude 72°30′ S, longitude 156° E.* Their positions change continuously.

Lines of force may be mapped for the earth's magnetic field, as in Fig. 335. If a compass needle is delicately mounted so as to be free to turn about both a horizontal and a vertical axis (Fig. 336), it will place itself tangent to the line of force at that point. Such a needle will be found to point neither truly north nor nearly horizontal.

FIG. 336. Free Magnetic Needle

A plane passed through the needle and the center of the earth is called the **magnetic meridian.**

A plane which passes through the axis of the earth and the point of observation (center of the needle) is called the **true meridian.**

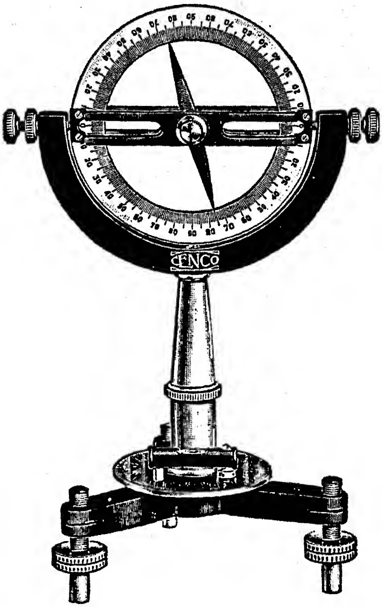


FIG. 337. Combined Compass and Dip Needle. (Courtesy Central Scientific Co.)

The angle which the magnetic meridian makes with the true meridian is called the **declination,** or **deviation,** of the compass. That is, declination is the angle between the direction of north as indicated by the compass needle and the direction of true north.

The angle which a suspended magnet makes with the horizontal plane through its center is called the **dip, or inclination.** Declination and dip are usually determined with different instruments: for the former the needle turns in a horizontal plane only; for the latter a "dipping needle," turning only on a horizontal axis, is used. An instrument that may be used for both purposes is shown in Fig. 337. Lines of equal declination (isogonic lines) for the

* Courtesy of the U. S. Coast and Geodetic Survey.

United States are sketched in Fig. 338 for each 5°, approximately correct for 1932.

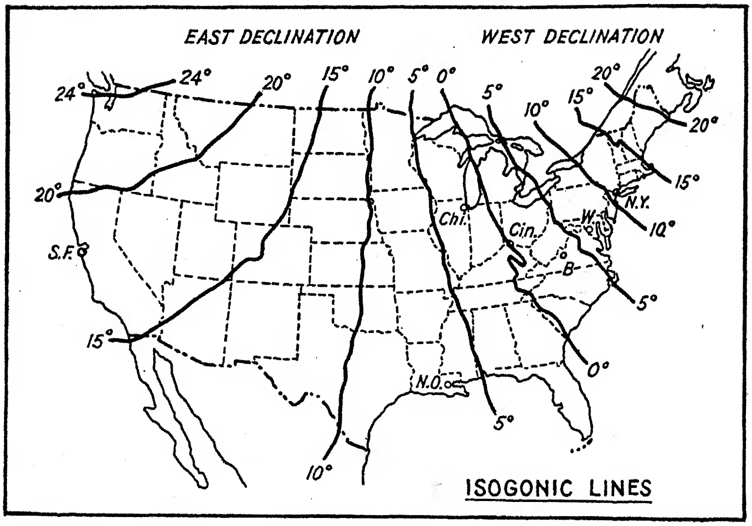


FIG. 338. Isogonic Lines

In order to determine the intensity of the earth's magnetic field, its horizontal component H_h is first determined by means of a magnet suspended to oscillate like a torsion pendulum.* The angle of dip θ being known, the total intensity H_t is then found by the following equation, based on Fig. 339a:

FIG. 339a. Total Intensity

$$H_t = \frac{H_h}{\cos\theta}. \quad (294)$$

The ratio of the horizontal and vertical components of the earth's magnetic field may be determined by means of the earth inductor, Fig. 339b.

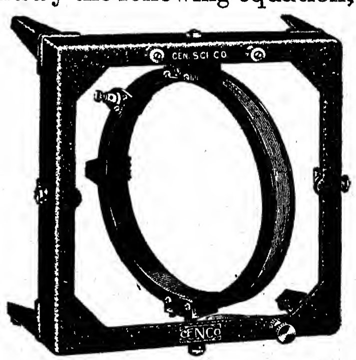


FIG. 339b. Earth Inductor. (Courtesy Central Scientific Co.)

The intensity, declination, and dip vary throughout the year (secular), throughout the day (diurnal), and on

* J. J. Thomson, *Electricity and Magnetism* (Cambridge University Press, 1921), p. 160.

occasion (magnetic storms). Secular variations follow closely the 11-year cycle of sunspot change, with which they are undoubtedly connected. Diurnal variations are due to the varying effects of the sun and moon during the day and night. Magnetic storms are indicated by large changes in the direction and intensity of the earth's field. They appear to be associated with sunspot activity, and frequently accompany displays of the aurora borealis.

Terrestrial magnetism has not yet been satisfactorily explained.

388. Theories of magnetism. As has already been mentioned, if a magnet is broken, a new north and a new south pole appear at the sides of the break. This is true regardless of how often the successive pieces are broken. Hence, if we could continue the breaking-up process until the magnet was reduced to its molecules, we might expect each molecule to exhibit a north and a south pole; and if the molecules are all alike, all the north poles should be equal and all the south poles equal. Also, since we find magnetizable material in neutral pieces of all sizes, the simplest assumption is that the individual north and south poles are equal and hence neutralize each other when arranged at random.

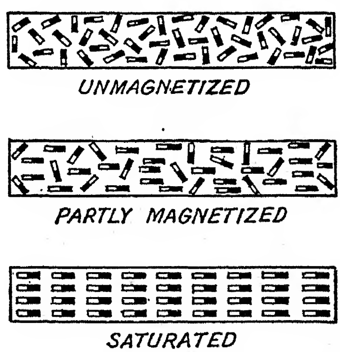


Fig. 340. Molecular Theory of Magnetization

From such considerations we have Ewing's theory: The molecules of all magnetizable substances are tiny magnets, each having a north and an equal south pole. In the unmagnetized state, these elementary magnets are arranged in groups so that the unlike poles neutralize one another (Fig. 340). On being magnetized, these elementary magnets are oriented so that a majority of norths point in one direction, and of souths in the opposite direction; and the material exhibits poles.

At first, in the magnetizing process, only the less stable groups will break up. But as magnetization proceeds, the poles become stronger, until finally no amount of stroking or increase in magnetizing current causes any further increase in the pole strength. The material is then said to be **saturated,** and we should expect

this when all the molecules are turned with like poles in one direction.

While this theory merely pushes the responsibility for magnetism back to the molecule, on this basis it explains very satisfactorily the following facts:

1. Equality of poles.
2. Appearance of new poles at a break.
3. Induced pole unlike inducing pole.
4. Gradual magnetization.
5. Saturation.
6. Loss of poles at high temperatures.

In an effort to explain why the molecules of a substance should have magnetic poles, André Marie Ampère suggested as early as 1825 that in each molecule of a magnetizable material there is a circular current of electricity ("Amperian current"), so that each such particle is in reality a tiny electromagnet (Sec. 455). Modern theory indicates that Ampère made a marvelously good guess.

On the basis of the Rutherford-Bohr theory of atomic structure every atom consists of a positive nucleus around which electrons revolve in various orbits. More recently, Goudsmit has shown that the fine structure of spectral lines is evidence that the electrons are also spinning about their own axes.

Rowland's experiment (Sec. 408) shows that any charge of electricity in motion is an electric current and produces a magnetic field. Hence Ampère's hypothetical currents are actually accounted for by our present knowledge of atomic structure, but inhere in the atom rather than in the molecule.

These whirling charges give to each atom a definite magnetic moment, and associated with it is a definite angular momentum. We should therefore expect the atoms to behave like little gyroscopes and to conform to the laws of quantum mechanics. If the former prediction is true, on spinning a piece of iron at high speed the tiny gyroscopes should set themselves with their axes parallel to the axis of rotation; and hence the specimen should exhibit magnetic poles due to the rotation alone. This experiment has been made by S. J. Barnett, and feeble poles were found as expected. On the basis of the quantum theory, J. C. Slater has accounted for the exceptional magnetic properties of iron, cobalt, and nickel.

Diamagnetism is the magnetic property best understood. It is thought to be due to precession of the little gyroscopes about axes parallel to the direction of the field. It appears to be exhibited by all substances; but, being a small effect, it is often masked by paramagnetism or ferromagnetism.

Paramagnetism is accounted for by partial orientation of the atomic magnets ("magnetons") of substances in which there is an unbalanced magnetic moment of the atom.

Ferromagnetism is least well understood. It seems necessary to suppose that in ferromagnetic substances some 10^{15} atoms are associated together in a **"domain"** which has a large resultant magnetic moment, and that these domains are readily aligned by a magnetic field. While such domains have not been definitely identified in the structure of the material, on alignment they produce a click that may easily be amplified and heard.

There is much experimental evidence in support of these theories.*

PROBLEMS

1. Two magnetic poles, one of 40 and the other of 15 unit poles, are placed 5 cm apart in air. Find the force with which the poles repel each other.

2. Find the field intensity at a point 5 cm from the north pole and 12 cm from the south pole of a magnet if the poles are 13 cm apart and have a pole strength of 60 emu.

3. Two magnetic poles, one of 60 and one of 25 unit poles, are placed 8 cm apart in air. Find the force with which the poles repel each other.

4. Two magnetic poles of 50 and 15 unit poles, respectively, are 10 cm apart in air. Find the force with which the poles repel each other.

5. The two poles of a bar magnet are 6 cm apart and of 300 unit poles strength. What force does the magnet exert upon a pole of 15 units placed on the perpendicular bisector of its length and 4 cm above its center line?

6. Calculate the moment of force exerted on a magnet 4 cm long if the axis of the magnet makes an angle of 30° with a field of intensity 0.18 oersted. The strength of each magnet pole is 200 unit poles.

7. What is the torque exerted on a magnet 6 cm long if the axis of the magnet makes an angle of 60° with a magnetic field of 0.22 oersted intensity? The strength of each pole is 100 unit poles.

* References:

S. R. Williams, *Magnetic Phenomena* (New York, McGraw-Hill Book Co., 1931).

S. Dushman, "Theories of Magnetism," *General Electric Review* (May, Aug., Sept., Oct., Dec., 1916).

K. K. Darrow, Theory of Magnetism, *Bell Technical Journal* (April 1936).

F. Bitter, *Introduction to Ferromagnetism* (New York, McGraw-Hill Book Co., 1937).

8. Find the field intensity at a point 3 cm from the north pole and 4 cm from the south pole of a magnet if the poles are 5 cm apart and the pole strength is 40 unit poles.

9. Find the field intensity at a point 5 cm from the north pole and 12 cm from the south pole of a magnet if the poles are 13 cm apart and the pole strength is 50 emu.

10. What is the total flux from a pole whose strength is 75 emu?

11. The horizontal component of the earth's field is 0.22 oersted and the resultant is inclined 21° to the vertical. Find the resultant strength of the field and the vertical component.

12. The resultant field intensity due to the earth at a certain point is 0.56 oersted and is inclined 30° to the vertical. Find the horizontal and vertical intensities.

13. How many lines of force pass through 1 m^2 of floor area where the total strength of the earth's magnetic field is 0.5 oersted and the dip is 69°?

14. How many lines of force pass through 1 ft^2 of floor area where the total strength of the earth's magnetic field is 0.4 oersted and the dip is 72°?

15. How many lines of force pass through 1 m^2 of a vertical wall where the total strength of the earth's field is 0.25 oersted, the declination 8° west, and the dip 70°: (a) when the wall runs north and south? (b) when the wall runs east and west?

ELECTRODYNAMICS

I cannot help thinking while I dwell upon them that this discovery of magnet-electricity is the greatest experimental result ever obtained by an investigator.

—John Tyndall

CHAPTER XXVI

FUNDAMENTAL LAWS AND INSTRUMENTS

389. Electric current. After Stephen Gray discovered in 1730 that metals could be electrified like other bodies if held by an insulating handle, it was generally assumed that the reason they could not be electrified without such handles was because electricity flowed through conductors, more or less as water flows through a pipe, and escaped through the body to the ground. We are now reasonably sure from the work of Sir J. J. Thomson and R. A. Millikan that **a current of electricity is a stream of electrons.** The "free electrons" in a conductor are repelled from the negative end of the conductor and drift toward its positive end.

Quantitatively, **electric current I is defined as the time rate of flow of electricity.** Average current (I) is expressed algebraically by

$$I \equiv \frac{Q}{t} \tag{295}$$

where Q is the quantity of electricity that flows past a given point in the time t.

Electrodynamics deals with electricity in motion. Since electricity in motion is always accompanied by a magnetic field, electrodynamics is also called electromagnetics, or current electricity.

390. The voltaic cell. The electrostatic generators described in Secs. 369–371 are not satisfactory for maintaining a current of electricity, for reasons that will be understood as we proceed. The first device for this purpose was the **simple cell** of Alessandro Volta who in 1799, while endeavoring to explain the famous frog leg experiment of Galvani, found that when a strip of copper and one of zinc were dipped into dilute acid and their outer ends connected by wires to the nerve and muscle, the frog legs twitched as when excited from the static machine.

That static charges actually appear on the copper and zinc strips may be shown by the following experiment. Let a con-

denser C (Fig. 341) be supported on an insulating stand and have its upper plate provided with an insulating handle, the plates being separated by a thin, unbroken layer of shellac or paraffined paper. Let the upper plate be connected to the ground and the lower plate to a gold leaf electroscope J. If the copper strip A of the simple cell is then connected to the lower plate of the condenser, and the zinc strip B to the upper plate, the electroscope will give no indication, not being sufficiently sensitive.

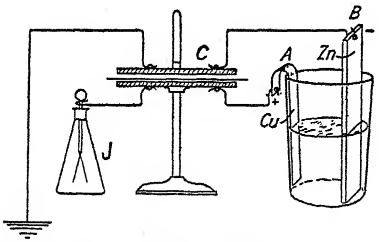


FIG. 341. Polarity of Simple Cell

But let the connection to the cell be broken with well-insulated pliers, and the upper plate lifted gradually by means of its insulating handle. Then, as the plate is raised, the leaves of the electroscope will diverge, because the capacitance of the condenser decreases as the distance between its plates increases, and the charge being constant, the potential difference between its plates increases in proportion to the work done in separating the charges on those plates.

If the charge on the electroscope is tested, it will be found positive. Hence the **copper** terminal is called the **positive pole** and the **zinc,** the **negative pole,** of the simple cell. (Zinc is negative in all ordinary cells.)

The copper and the zinc were at the same potential before they were placed in the acid. As a result of the chemical action, equal + and − charges are segregated upon A and B, respectively; and since positive work would have to be done on each + charge to move it from B to A, the potential of A is higher than that of B, according to the definition of Sec. 353. Consequently, if A and B are connected by a conductor, + charges should flow along it from A to B, and in the experiment this appears to be true. Hence it became *conventional* to say that *positive electricity flows from the positive to the negative terminal of a cell, or of any generator.*

As a matter of fact, we are now quite certain that exactly the reverse of this takes place. **Electrons flow from the negative to the positive terminal along the wire.** The discrepancy is due to the fact that we arbitrarily chose the unit + charge as our test charge (Sec. 343). Had the unit − charge been chosen as the test

charge, our mathematics would have been in exact accord with the facts. However, this difficulty, being due to a mere convention of signs, does not affect the final results of an experiment, and will be corrected gradually as the literature of science increases.

The simple Voltaic cell will maintain a current of electricity for a considerable time in a wire connecting its terminals, as may be shown by any of the current-detecting devices described in the following paragraphs. The defects of this cell and the improved forms intended to eliminate these defects are described in Chap. XXIX.

FIG. 342

The conventional representation of a primary cell is shown in Fig. 342, where the long thin line represents the positive plate and the short thick one, the negative plate.

Several cells used in conjunction are called a **battery.**

391. Oersted's experiment. In order to measure currents of electricity, we must consider next a phenomenon accidentally discovered in 1819 by Hans Christian Oersted of the University of Copenhagen, while experimenting before his class. A wire was placed parallel to a magnetic needle, as in Fig. 343. On producing a current in the wire by means of a battery (several cells), he noted that the needle was deflected strongly toward a position at right angles to the wire. On reversing the current, the deflection was reversed.

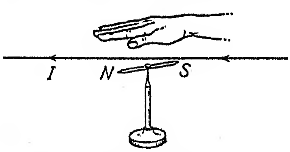


FIG. 343. Oersted's Law

The behavior of the needle may be predicted by **Oersted's law.** If a wire is parallel to a magnetic needle and a current of electricity is produced in the wire, the needle will deflect according to the following rule: **Place the right hand so that the wire lies between the palm of the hand and the needle. Then, if the fingers point in the direction of the current, the thumb will point to the side to which the north pole will be deflected.**

This experiment was the discovery of electric motor action, the principle upon which the numberless electric motors now made all operate. It was the first intimation that **there is always a magnetic field associated with electricity in motion.**

392. Magnetic field due to a current in a straight wire. The behavior of the magnetic needle in Oersted's experiment is readily

understood if we plot the magnetic field about a straight wire on a card through which the wire passes perpendicularly. This may be done either with iron filings or by means of a small compass.

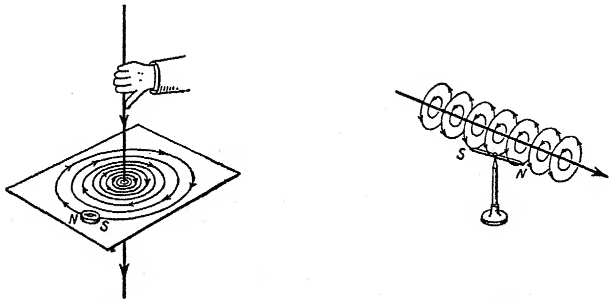


FIG. 344. Field about a Straight Wire

By either method, the lines of force are found to be concentric circles with the wire at the center (Fig. 344).

The direction of these magnetic lines is given by the following rule: **Grasp the wire with the right hand, the thumb pointing in the direction of the current. The fingers will then point in the direction of the magnetic lines of force.**

393. Magnetic field of current in a helix or solenoid. If the wire of the preceding paragraph is bent into a helix and the lines are traced out, it will be found that the magnetic fields due to the current in the individual turns of the helix give a resultant field, as shown in Fig. 345, which is similar to the field of a bar magnet (Fig. 332). We might expect, therefore, that the end of the coil from which the lines emerge would behave like a north pole; and the end where they enter, like a south pole. This is found to be the fact: a helical or a circular current has all the properties of a magnet.

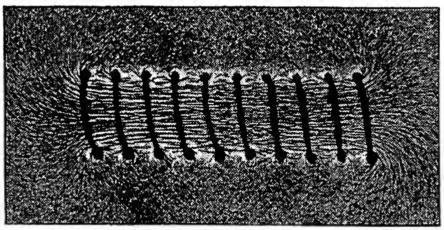


FIG. 345. Magnetic Field in Solenoid

The polarity of a helix, or solenoid, is readily determined by the following rule: **Grasp the helix with the right hand with the fingers**

pointing in the direction of the current. The thumb will then point to the north end of the coil (Fig. 346).

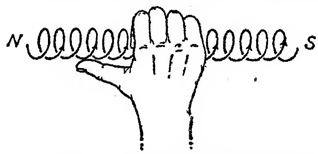


Fig. 346. The Right Hand Rule for the Helix

To determine whether one is looking at the north or the south end of a coil, write an N or an S in the end view of the coil. If the current has the direction of the arrow points of the N, it is the north end of the coil; otherwise the south (Fig. 347).

Fig. 347. Polarity of Coils Seen End-On

394. Ampère's law. Since all motion is relative, we should expect that if the magnet in Oersted's experiment were held fixed, the wire in which the electricity flows would move. Ampère (1775–1836), professor of physics at the Polytechnic School in Paris, investigated the relations between electric currents and magnetic fields in a brilliant manner. Somewhat simplified, his law may be stated as follows:

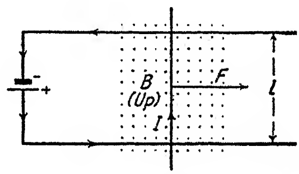


Fig. 348. Ampère's Law

If a length l of wire containing a current I lies in and perpendicular to a magnetic field of flux density B, it is acted upon by a force F which is proportional to B, I, and l and which is perpendicular to B and I (Fig. 348). Algebraically,

$$F \propto BIl$$

$$F = kBIl \tag{296}$$

where the constant of proportionality k depends upon the choice of units. If the wire is free to move, the direction of its motion will be the same as the direction of F.

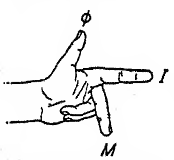


Fig. 349. Fleming's Motor Rule

The direction of the force F is conveniently determined by **Fleming's motor rule:** Set the first finger (thumb), the index finger, and the middle finger of the *left* hand at right angles to one another. Then, if the first finger is in the direction of the flux, and the index finger in the direction of the current I, the middle finger will be in the direction of the motion (Fig. 349). *This is the only left hand rule.*

Fleming's rule is easily remembered from the mnemonic scheme that the initial letters of corresponding things are the same:

*F*irst finger —*F*lux
*I*ndex finger —*I* (current)
*M*iddle finger—*M*otion.

Thus far no unit of current has been defined. Since k in Ampère's law depends upon the choice of units, it will be simplest to choose a unit of I so as to make k equal to unity. Accordingly, **the abampere is defined as that current which, in a wire one centimeter long in and at right angles to a magnetic field of one line per square centimeter, is acted upon by a force of one dyne at right angles to the current and the flux.**

If we put the values from this definition back into Eq. (296) above, we have:

$$1 \text{ dyne} = k \times 1 \frac{\text{line}}{\text{cm}^2} \times 1 \text{ abampere} \times 1 \text{ cm}$$

whence

$$k = 1, \text{ a pure number.}$$

Hence, using these *absolute electromagnetic units* (emu), we have Ampère's law:

$$F = BIl \quad \text{provided emu are used.} \tag{297}$$

395. Units of electric current. The abampere, as defined above, is the fundamental unit of electric current upon which all others are based. It is rather large for most purposes; therefore a **practical unit** of current, the ampere, is defined.

The ampere is one-tenth of an abampere:

$$1 \text{ ampere} \equiv \frac{1}{10} \text{ abampere.} \tag{298}$$

By using with this the metric prefixes, we get the smaller units:

$$1 \text{ milliampere} \equiv \frac{1}{1000} \text{ ampere}$$

$$1 \text{ microampere} \equiv \frac{1}{1{,}000{,}000} \text{ ampere.}$$

396. Units of quantity of electricity. In electrostatics the unit of quantity of electricity, the statcoulomb, was defined in order to simplify Coulomb's law. No unit of current was defined in electrostatics, because that would have involved the incongruity of considering electricity in motion when studying electricity at rest. We may now define a unit of current, the statampere, as the current in a conductor when one statcoulomb of electricity passes any given cross section of the conductor in one second. But this unit is not used in electrodynamics because it does not make the constant k equal unity in Ampère's law, as does the abampere.

Consequently, in electrodynamics we take the abampere as the basic unit and define the **abcoulomb** as the quantity of electricity that passes any given cross section of a conductor when a current of one abampere is maintained for one second.

Clearing fractions in Eq. (295),

$$Q = It$$

so that

$$1 \text{ abcoulomb} \equiv 1 \text{ abampere} \times 1 \text{ sec} \tag{299}$$

and

$$1 \text{ coulomb} \equiv \tfrac{1}{10} \text{ abcoulomb} \tag{300}$$

from which

$$1 \text{ coulomb} = 1 \text{ amp} \times 1 \text{ sec.} \tag{301}$$

397. Galvanometers are instruments for measuring electric current. The facts of the foregoing paragraphs provide a basis for the construction of current measuring instruments. J. S. C. Schweigger made the first galvanometer in 1820 by winding the wire in Oersted's experiment many times around the magnetic needle. In 1837 Pouillet developed this arrangement of stationary coil and moving magnet into the tangent galvanometer, in which the current is proportional to the tangent of the angle of deflection. It is now seldom used because of the necessity of orienting its coil in the plane of the earth's field.

The **moving coil galvanometer** has a stationary magnet and a moving coil. This permits the use of a magnet strong enough to make the earth's field negligible. It was invented independently

by Desprez and D'Arsonval about 1862, but is generally known as the D'Arsonval galvanometer (Fig. 350).

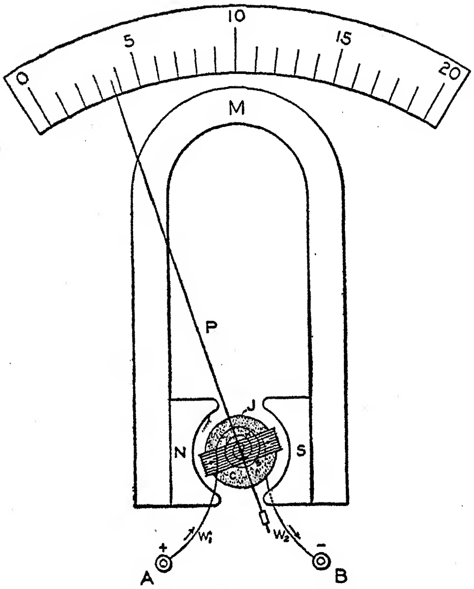


FIG. 350. Moving Coil Galvanometer

The rectangular moving coil C rotates about a vertical axis through its center and carries the pointer P, or a mirror, rigidly fixed to it.

Electricity flows to C by means of the light spiral spring W_1, passes around the coil as shown by the arrow, and leaves by the spiral spring W_2. This makes the upper face of C a north pole and the lower, a south. These poles of the coil are repelled by the like poles of the magnet M, causing the whole moving system to rotate clockwise.

A cylindrical soft-iron core J nearly fills the rectangular opening in C but does not turn with the coil. If this core and the pole pieces of M are properly shaped, the coil will turn in a uniform, radial field. It is then easily shown that

$$I = k\theta \tag{302}$$

where θ is the angle of deflection of the moving system and k is a constant characteristic of the instrument.

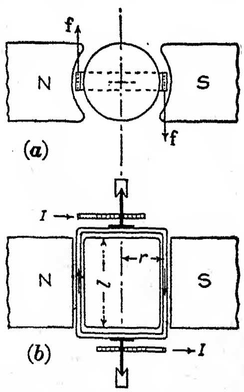


FIG. 351

Figures 351a and 351b show the horizontal and vertical cross sections, respectively, of the coil and magnet of the galvanometer of Fig. 350. The coil is rectangular. Let the length of each side that is perpendicular to the flux be l cm, and the current in the coil be I abamperes.

Then, by Ampère's law, the force on each conductor is:

$$f = BlI \text{ dynes}$$

and the torque due to this force is

$$\mathfrak{T}_1 = fr = BlIr \text{ cm-dynes.}$$

If there are n turns on the coil, there are $2n$ of these conductors, half on each side of the center, all producing clockwise torques, as shown.

The total deflecting torque on the coil is therefore:

$$\mathfrak{T}_D = 2nBlIr \text{ cm-dynes.} \qquad \text{(a)}$$

As the coil deflects, the springs are twisted and tend to untwist with a restoring torque whose value is proportional to the angle of twist θ, according to Hooke's law. That is, the restoring torque is:

$$\mathfrak{T}_R = c\theta \quad \text{cm-dynes} \qquad \text{(b)}$$

where c is the torsional constant.

The coil will obviously take up a position in which the deflecting torque is balanced by, or equal to, the restoring torque. Therefore, from Eqs. (a) and (b),

$$2nBlIr = c\theta$$

$$I = \frac{c}{2nBlr}\theta.$$

But $\dfrac{c}{2nBlr}$ is a constant, say, k.

Therefore, $$I = k\theta. \qquad (303)$$

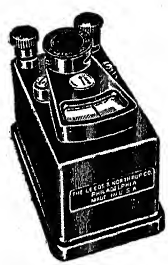
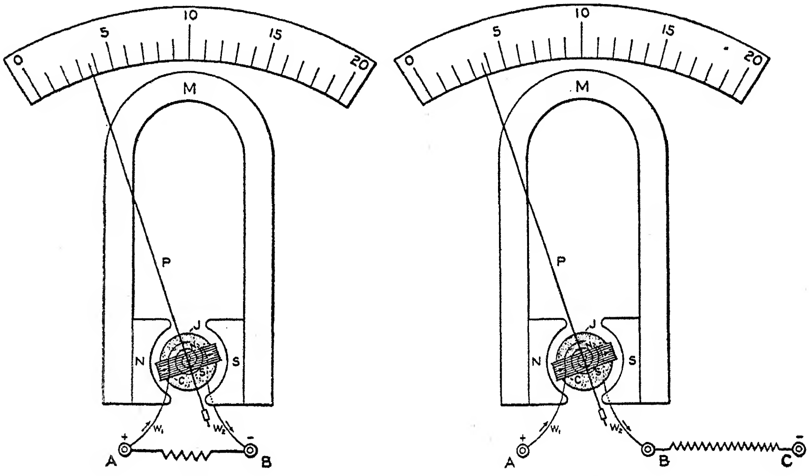


FIG. 351(*c*). Portable Galvanometer. (Courtesy Leeds & Northrup Co.)

Hence, in a properly constructed moving coil galvanometer, the current is proportional to the angle of deflection, and the instrument will therefore have an equally divided scale. A portable galvanometer is shown in Fig. 351c.

398. Ammeter and voltmeter. An **ammeter** is a galvanometer having a conductor of low resistance (a **shunt**) connected **in parallel** with its moving coil and its scale calibrated in amperes (Fig. 352a).

A **voltmeter** is a galvanometer having a conductor of high resistance (a **multiplier**) connected **in series** with its moving coil, and its scale calibrated in volts (Fig. 352b).



(*a*) (*b*)

Fig. 352. Internal Connections of (*a*) Ammeter, (*b*) Voltmeter

399. Connections of ammeter and voltmeter to a circuit. The above-mentioned connections of the shunt and the multiplier are generally permanent *internal* connections made at the factory. The connections of these instruments to a circuit, i.e., the **external** connections to the binding posts, should be made as follows.

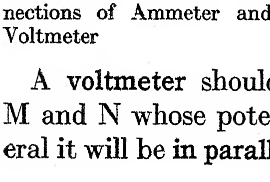


Fig. 353. External Connections of Ammeter and Voltmeter

An **ammeter** should always be connected **in series** in the circuit in which the current is to be measured (Fig. 353).

A **voltmeter** should always be connected to the two points M and N whose potential difference is to be measured. In general it will be **in parallel** with a portion of a circuit, as in Fig. 353.

400. The ballistic galvanometer. A ballistic galvanometer is a galvanometer used to measure the **quantity** of electricity suddenly discharged through it.

In Sec. 397 it was shown that when a steady current I is maintained in a moving coil galvanometer of proper design, the deflection θ is proportional to the current.

If, however, a quantity Q of electricity is **suddenly** passed through a gal-

vanometer whose moving system has sufficient inertia, the throw θ, or maximum deflection, may be shown * to be proportional to the quantity Q.

That is,

$$Q \propto \theta$$

or

$$Q = k'\theta \tag{304}$$

where k' is a constant of proportionality characteristic of the particular galvanometer. The numerical value of k' may be determined by discharging through the galvanometer the charge from a condenser of known capacitance that has been charged to a known potential difference, such as that of a standard cell.

Any galvanometer may be used ballistically, but in order to be accurate the moment of inertia of the moving system must be such that it acquires its maximum angular velocity while the deflection is still negligibly small.

401. Electromotive force. When electricity flows in a circuit (Fig. 354) from any point 1 of higher potential to any point 2 of lower potential, it does work, such as turning a motor or merely heating the wires of the circuit itself. Points 1 and 2 would quickly come to the same potential, and the flow would cease, unless some device were provided to maintain the potential difference between them. Such a device is called a source of electromotive force, or generator, in this case the cell CZ.

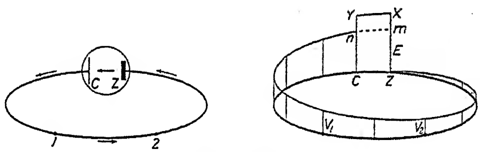


FIG. 354.

Each unit quantity of electricity in flowing from 1 to 2 does an amount of work equal to the fall of potential between 1 and 2; and the total work done by a unit quantity when it flows completely around the circuit is the summation of all these potential differences, or potential drops, from point to point around the circuit. To supply this energy, the generator must do an equal amount of work on each unit quantity as it passes through the generator, thereby maintaining the potential differences around the circuit.

The electromotive force (emf) of a generator is the work done by the generator on each unit quantity of electricity that passes through it. If the generator is in a simple series circuit, its emf

* Gilbert, *Electricity and Magnetism*, *op. cit.*, p. 170.

is equal the work required to move a unit quantity of electricity completely around the circuit.

Algebraically,
$$E \equiv \frac{W}{Q} \tag{305}$$

where E is the emf of a generator which does an amount of work W on Q units of quantity of electricity as they pass through it.

If a generator is producing a current, some of its emf is required to do the work of forcing the electricity through the generator's own internal resistance; the difference of potential between the terminals of the generator is therefore less than its emf. The emf of a generator equals the potential difference between its terminals *only* when it is not producing a current; and the measurement must be made with an instrument that requires no current, such as an electrostatic voltmeter or a potentiometer.

On clearing fractions above, we have the very important equation:

$$W = EQ = EIt \tag{306}$$

which we shall take up again in Sec. 418.

402. Law of Henry and Faraday. Although the Voltaic cell was the first and for many years the only source of electric current, today electromagnetic generators are used for all electric power purposes greater than that of a radio set. The discovery that an electromotive force could be developed by electromagnetic induction was made independently by Joseph Henry in America (1830) and Michael Faraday in England (1831).

The law may be stated in two ways which are equivalent:

1. **Whenever the magnetic flux through a coil changes, an emf E is induced in the coil proportional to the number of turns N on the coil and to the rate of change of flux through the coil.**

Algebraically,

$$E \propto N\frac{(\Phi_2 - \Phi_1)}{t} \tag{a}$$

where Φ_1 is the initial flux through the coil;
Φ_2 is the final flux through the coil; and
t is the time required for the flux to change uniformly from Φ_1 to Φ_2.

Thus, let the coil of Fig. 355 be moving from left to right through a non-uniform magnetic field, as shown. Then the flux through the coil is increasing, and an emf is induced in it which will produce a current as shown by the arrows if the circuit is closed between A and B. If the field were uniform, the total flux through the coil would not be changing and the emf induced in the coil as a whole would be zero.

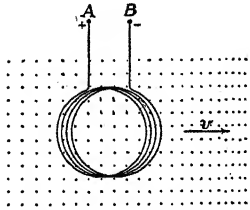


FIG. 355. Electromotive Force Is Induced in Coil

This law is often stated as follows:

2. Whenever a conductor cuts across lines of magnetic flux, an electromotive force is induced in it which is proportional to the rate of cutting the lines of flux. Emf so induced is called "motional emf."

If there are N conductors cutting at the same rate and so connected that the emf's all assist one another, then the total emf E is as above:

$$E \propto N\frac{(\Phi_2 - \Phi_1)}{t} \tag{a}$$

where $(\Phi_2 - \Phi_1)$ is the number of lines of flux cut by each conductor; t is the time each conductor takes to cut all the lines; and the rate of cutting must be uniform.

When an emf is induced in a conductor, if the circuit is closed, electrons will flow around the circuit. This is equivalent to the conventional flow of positive electricity the other way around. The **sense** of flow of positive electricity (clockwise or counterclockwise, right or left, up or down) is commonly called the **direction** of the emf that produces the current. Strictly speaking, emf is a scalar quantity and therefore has no direction. But scalars have sense (sign), and what is called direction of emf is actually the sense in which the scalar emf increases along the conductor in which it is induced.

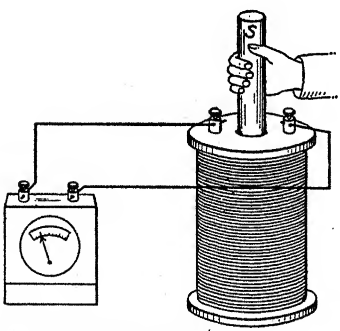


FIG. 356. Electromagnetic Induction

Thus, when the N-pole of the magnet of Fig. 356 is thrust into

the coil, an emf is induced that causes a current which deflects the galvanometer G to the right, say. When the magnet is withdrawn, G is deflected to the left. If the circuit is now opened somewhere and a cell is inserted so as to produce a current that will deflect G to the right, it will be found that this current is in such a sense around the coil as to make its top end a north pole. That is, the current due to the induced emf as the magnet enters opposes its entrance. By tracing it out, one will find that the current produced on withdrawing the magnet opposes the withdrawal.

On reflection, it is clear that the induced polarity of the coil must be like that of the approaching pole. For otherwise one would need only to start a magnet pole into a coil; then, if an opposite pole were induced, the magnet would be drawn in, inducing more emf without work. This would provide an unlimited source of energy without doing any work, which is contrary to our cherished belief in conservation of energy.

These facts are summed up in **Lenz's law: An induced emf has always such a sense as to oppose the change that causes it.**

This law is represented algebraically by making the second member of Eq. (a) negative:

$$E \propto -N\frac{(\Phi_2 - \Phi_1)}{t} \tag{b}$$

$$E = -kN\frac{(\Phi_2 - \Phi_1)}{t} \tag{c}$$

where k depends upon the units chosen for the other factors.

Since we have chosen no electromagnetic unit of potential difference, or emf, we now choose one so that k will be unity:

The abvolt is defined as the electromotive force induced in a coil of one turn when the flux through it changes at the rate of one line per second; or, it is the potential difference induced between the ends of a conductor which cuts across lines of flux at the rate of one line per second.

On writing these units in Eq. (c), we have:

$$1 \text{ abvolt} = -k \times 1 \text{ turn} \times 1 \frac{\text{line}}{\text{sec}}$$

so that $k = 1$, as was intended.

The law of Henry and Faraday then becomes:

$$E = -N\frac{(\Phi_2 - \Phi_1)}{t} \tag{307}$$

provided that the following electromagnetic units (emu) are used:

E is in abvolts;

N is the number of turns, or the number of conductors;

$\Phi_2 - \Phi_1$ is the total change of magnetic flux, or the total flux cut by each conductor; and

t is the time in seconds.

If the change or cutting of flux is not uniform, the law of Henry and Faraday must be written in the differential symbols:

$$E = -N\frac{d\Phi}{dt} \tag{308}$$

where $d\Phi$ is the change of flux in the time dt, which is taken so small that during each dt the rate of change may be considered constant.

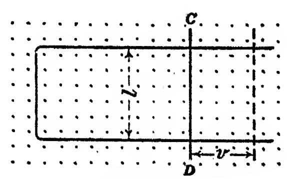


FIG. 357

For the emf induced in a conductor cutting across magnetic flux, a useful expression may be derived as follows (Fig. 357).

Let a bare wire CD of length l cm slide along two parallel bare conductors forming a loop as shown, with a velocity v cm/sec perpendicular to a uniform magnetic field of B lines/cm^2.

Then the area swept over in 1 sec is lv cm^2, and the flux cut in 1 sec is Blv lines. That is,

$$\frac{d\Phi}{dt} = Blv$$

so that

$$E = -N\frac{d\Phi}{dt} = -NBlv \tag{309}$$

where N is the number of conductors cutting as above and connected so that their emf's are cumulative.

It will be noticed that the rate of cutting lines by CD is the same as the rate of change of flux through the loop.

The law of Henry and Faraday is the **law of generator action.** A convenient method much employed by engineers for determin-

ing the direction of flow of electricity (direction of the current) due to an induced emf is the following.

Fleming's generator rule.* Set the first finger (thumb), the index finger, and the middle finger of the **right hand** at right angles to one another. Then, if the first finger is in the direction of the flux, and the middle finger in the direction of the motion, the index finger will be in the direction of the current I produced by the emf.

403. Units of potential and electromotive force. The abvolt, as defined in the preceding section, is far too small for ordinary purposes. Hence we define the *practical unit* of potential and emf, the *volt*, as one hundred million abvolts. Algebraically,

$$\left.\begin{aligned} 1 \text{ volt} &\equiv 10^8 \text{ abvolts} \\ 1 \text{ millivolt} &\equiv \frac{1}{1000} \text{ volt} \\ 1 \text{ microvolt} &\equiv \frac{1}{1{,}000{,}000} \text{ volt.} \end{aligned}\right\} \tag{310}$$

At this point the question will naturally arise in one's mind: Why do we not continue to use the unit of potential defined in electrostatics? On reflection it will be seen that the question is already answered: The electrostatic unit of quantity would not make the proportionality factor unity in the law of Henry and Faraday. We shall presently be able to deduce the relation:

$$1 \text{ statvolt} = 300 \text{ volts.}$$

404. Ohm's law. We are accustomed to think of electricity as flowing from points of higher to points of lower potential, in much the same way that water flows from points of higher to points of lower pressure. Since greater difference of pressure causes a greater flow of water, we should expect that greater difference of potential would produce a greater flow of electricity, i.e., more current. Experiment shows that it does.

The relation of potential difference to the resulting current was investigated by Georg Simon Ohm in 1826. His conclusion constitutes *one of the most used laws* in the whole of electrical science.

* John Ambrose Fleming, *Electrician*, *XIV* (1885), 396.

Ohm's law: The current I in a metallic conductor is directly proportional to the difference of potential V between its terminals. Algebraically,

$$V \propto I.$$

Therefore,

$$V = RI. \tag{311}$$

The factor of proportionality R is called the *resistance* of the conductor and depends upon the material, the dimensions, and the temperature of the conductor, as we shall see in the next chapter.

The resistance to the flow of electrons through metallic conductors is believed to be due to their collisions with the atoms of the material, and it remains constant if the temperature of the conductor is constant. In electrolytic conductors or in vacuum tubes, this is not true; hence, in these cases, Ohm's law does not hold. Electrical resistance is somewhat like the resistance that water encounters in flowing through pipes, but here also the resistance is not constant, changing when the rate of flow, or current, changes. Hence a law similar to Ohm's would not hold true for water.

Ohm's law applies to a simple circuit as a whole, as well as to any part of it, provided we give the proper interpretation to V and to R. This will be taken up in Sec. 442.

405. Units of resistance. Solving Eq. (311),

$$R \equiv \frac{V}{I}. \tag{312}$$

Here R will be unity when V and I are both unity; i.e.,

$$1 \text{ abohm} \equiv \frac{1 \text{ abvolt}}{1 \text{ abampere}}. \tag{313}$$

Hence, in the electromagnetic system: **The abohm is defined as the resistance of a conductor in which a current of one abampere is maintained by a potential difference of one abvolt.** Similarly, in the practical system:

$$1 \text{ ohm} \equiv \frac{1 \text{ volt}}{1 \text{ ampere}}. \tag{314}$$

From these definitions we may deduce the relation between the ohm and the abohm, as follows:

By definition,

$$1 \text{ ohm} \equiv \frac{1 \text{ volt}}{1 \text{ ampere}}$$

by Eq. (310) $$1 \text{ volt} \equiv 10^8 \text{ abvolts}$$

by Eq. (298) $$1 \text{ ampere} \equiv 10^{-1} \text{ abampere.}$$

Substituting in Eq. (314),

$$1 \text{ ohm} = \frac{10^8 \text{ abvolts}}{10^{-1} \text{ abampere}} = 10^9 \frac{1 \text{ abvolt}}{1 \text{ abampere}}.$$

But by Eq. (313),

$$\frac{1 \text{ abvolt}}{1 \text{ abampere}} \equiv 1 \text{ abohm.}$$

Therefore

$$1 \text{ ohm} = 10^9 \text{ abohms.} \tag{315}$$

406. Law of continuity of current. Another law attributed to both Ohm and Faraday is in two parts:

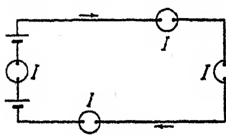


FIG. 358

1. **The electric current in an undivided circuit is the same at all sections of the circuit.** If we insert ammeters at any number of places in an undivided circuit, they will all read the same (Fig. 358).

2. **The sum of the currents in the branches of a divided circuit equals the current in the undivided part.**

In Fig. 359,

$$I = I_1 + I_2 + I_3. \,.$$

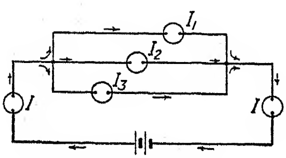


FIG. 359

407. Water analogues of electrical quantities. Since water is the fluid with which we are most familiar, it is helpful to have in mind the analogies between the various terms and practical units for electricity and for water. These are shown by the following table.

Electricity		Water	
Name	Unit	Name	Unit
Quantity	Coulomb	Quantity	Gallon
Current	Ampere	Current	Gal/sec
Potential	Volt	Pressure	$Lb/in.^2$
Resistance	Ohm	Friction	No Name

408. Rowland's convection experiment.* The question may well be raised whether what we are now calling an electric current is really electricity in motion, as was first suggested by Faraday. This question was answered satisfactorily by an experiment made at Berlin in 1876 by Henry A. Rowland of the Johns Hopkins University.

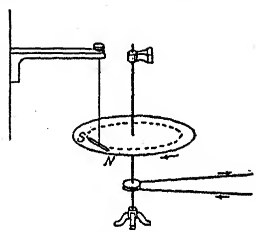


FIG. 360. Principle of Rowland's Experiment

Rowland mounted a disk of hard rubber about 21 cm in diameter so that it could be rotated about a vertical axis at high speed (Fig. 360).

When a static charge was placed upon the disk and the latter rotated at 3660 rpm, a magnetic needle suspended over the revolving charge was deflected exactly as it was in Oersted's experiment by what we called a current in a wire.

There remains little room for doubt that an electric current consists of electricity flowing through a conductor; that is, static electricity is the same kind as current electricity.

Franklin's kite experiment demonstrated that atmospheric electricity is the same as that developed by friction. There is now no reason to think that any kind of electricity exists except electrons, protons, and positrons.

PROBLEMS

1. A conductor 6 cm long is in and perpendicular to a magnetic field of 1000 lines/cm². It is urged out of the field by a force of 40 dynes. What is the current in the conductor in amperes?

2. A conductor 5 cm long is in and perpendicular to a magnetic field of 2000 lines/cm². It is urged out of the field by a force of 60 dynes. What is the current in the conductor in amperes?

* Henry A. Rowland, *Physical Papers* (Baltimore, The Johns Hopkins Press, 1902), p. 128.

3. A conductor 5 cm long, in and perpendicular to a magnetic field of 800 lines/cm^2 is urged out of the field by a force of 3.2 dynes. What is the current in the conductor?

4. There are 250 conductors each 12 in. long on the armature of a motor 16 in. in diameter. If the average flux density in which they stand is 50,000 lines/in.^{2}, what is the torque on the armature when the current in the conductors is 10 amp?

5. There are 300 conductors each 10 in. long on the armature of a motor 12 in. in diameter. If the average flux density in which they stand is 8000 lines/in.^{2}, what is the torque on the armature when the current in the conductors is 15 amp?

6. There are 250 conductors each 10 in. long on the armature of a motor 20 in. in diameter. If the average flux density in which they stand is 20,000 lines/in.^{2}, what is the torque on the armature when the current in the conductors is 12 amp?

7. A wire 20 cm long is moved through a uniform magnetic field having an intensity of 10,000 lines/cm^2. The wire moves perpendicularly to its own length and at right angles to the field. If the velocity of the wire is 1 m/sec, what emf is induced in the wire?

8. A coil of 200 turns is withdrawn from the middle of a bar magnet in 0.1 sec and the average emf induced is 0.2 volt. What is the pole strength of the magnet?

9. A coil of 400 turns is withdrawn from the middle of a bar magnet in 0.2 sec and the average emf induced is 0.4 volt. What is the pole strength of the magnet?

10. A wire 15 cm long is moved through a uniform magnetic field having an intensity of 6000 lines/cm^2. The wire moves perpendicularly to its own length and at right angles to the field. If the velocity of the wire is 2 m/sec, what emf is induced in the wire?

11. A train on a straight track runs north at a speed of 30 m/sec. If a car axle is 150 cm long and the vertical component of the earth's field is 0.2 gauss, what emf is induced in the axle and which end is at the higher potential?

12. If there is a current of 15 amp in a wire, how many electrons pass a cross section of the wire in 10 sec?

13. How many electrons per second pass through a 50-watt lamp if the current is 0.45 amp?

14. An electric soldering iron has a resistance of 24 ohms. What current will it take from 110-volt mains?

15. What is the resistance of a pressing iron in which there is a current of 5 amp when connected to 110-volt mains?

16. What potential difference is necessary to maintain a current of 12 amp in a resistance of 8 ohms?

17. The resistance of the moving coil of a galvanometer is 22 ohms and a current of 0.001 amp gives a full-scale deflection. If this galvanometer is to be used as a voltmeter having a full-scale reading of 100 volts, what must be the resistance of the multiplier?

18. The resistance of the moving coil of a galvanometer is 15 ohms and a current of 0.002 amp gives a full-scale deflection. It is to be made into an ammeter

to give a full-scale deflection for 10 amp. What should be the resistance of the shunt?

19. The resistance of the moving coil of a galvanometer is 15.3 ohms and a current of 0.0012 amp gives a full-scale deflection. If it is to be made into an ammeter to give a full-scale deflection for 10 amp, what should be the resistance of the shunt?

20. If the galvanometer of the preceding problem is to be used as a voltmeter having a full-scale deflection for 110 volts, what must be the resistance of the multiplier?

21. The resistance of the moving coil of a galvanometer is 14.6 ohms and a current of 0.0015 amp gives a full-scale deflection. If it is to be made into an ammeter to give a full-scale deflection for 15 amp, what should be the resistance of the shunt?

22. If the galvanometer of the preceding problem is to be used as a voltmeter having a full-scale deflection for 110 volts, what must be the resistance of the multiplier?

CHAPTER XXVII

RESISTANCE

409. Conduction in metals. As was mentioned in our discussion of electrostatics, some of the electrons associated with an atom of a metal are apparently quite free to move from atom to atom, whereas the reverse is true in nonconductors. But the protons, being strongly bound in the nuclei of the atoms, can be removed only with difficulty.

Consequently, when a potential difference is produced between two points of a conductor, electrons drift from the point of lower to the point of higher potential; and since electrons are negative, this is equivalent to a flow of positive electricity from points of higher to points of lower potential. Protons and positrons appear to take no part in metallic conduction.

When some water is pumped into one end of a long pipe that is already full, an equal amount of water is forced out at the other end at practically the same instant. Similarly, when some electrons are forced into one end of the Pacific cable, say, some other electrons are forced out at the other end at almost the same instant. This gives the appearance of enormous speed, whereas as a matter of fact the motion of the electrons may be relatively quite slow.

410. Resistivity. In their drift along the conductor, electrons will probably collide with one another and with the parts of the atoms that do not flow, and these collisions should impede the progress of the electrons. We should expect these "conductivity electrons" to encounter twice as much resistance in passing through a wire 20 cm long as in a wire 10 cm long, if both are of the same area A and of the same material. That is,

$$R \propto l \qquad \text{if } A \text{ is constant.}$$

Likewise, if two wires have the same material and length, but

one has half the area of the other, we should expect the smaller wire to offer twice the resistance of the larger; so that

$$R \propto \frac{1}{A} \qquad \text{if } l \text{ is constant}$$

or, in general,

$$R \propto \frac{l}{A} \qquad \text{when both } l \text{ and } A \text{ vary.}$$

This anticipated result is borne out by the most refined experiments.

Hence,

$$R = \rho \frac{l}{A} \tag{316}$$

where the factor of proportionality ρ depends upon the choice of units and upon the nature of the conductor. ρ is called the **resistivity,** or specific resistance, of the material. It is a characteristic property of the material, but varies with the temperature.

Resistivity is sometimes said to be **the resistance between the opposite faces of a centimeter cube** (not a cubic centimeter) **of the material.**

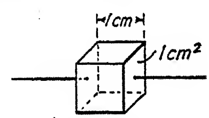


FIG. 361

This fact is easily found from Eq. (316), if we think of the conductor as being a cube of the material having each edge 1 cm long (Fig. 361). Then, for that conductor, $l = 1$ cm and $A = 1$ cm², so that Eq. (316) becomes:

$$R \text{ ohms} = \rho \frac{1 \text{ cm}}{1 \text{ cm}^2}$$

whence

$$\rho = R \text{ ohm-cm.} \tag{317}$$

Thus, the resistivity ρ is numerically equal to the resistance R between opposite faces of a centimeter cube.

For this reason, in many tables resistivities are stated in "ohms per (cm)^3" but the preferable unit is the **ohm-cm** as derived in Eq. (317).

In engineering it is customary to measure l in feet and A in circular mils. In that case the ρ is the resistance of a mil-foot of

the material; i.e., a wire 1 ft long and 1/1000 in. in diameter. (A mil is 1/1000 in. The area in circular mils is obtained by squaring the diameter in mils.)

Conductance is the reciprocal of resistance. Its unit is therefore the $(\text{ohm})^{-1}$; sometimes called the "reciprocal ohm," or "mho."

Conductivity is the reciprocal of resistivity. Its unit is the $(\text{ohm-cm})^{-1}$, or the "mho per cm."

411. Variation of resistance with temperature. According to the kinetic theory of matter, the velocities and the mean free path of the atoms of a body increase with temperature. Conductivity electrons participate in this atomic agitation, and we should therefore expect that the resistance of a substance would be affected by a change of temperature. The relation is found to be:

$$R_t = R_0(1 + \alpha_0 t + \beta_0 t^2 + \gamma_0 t^3) \tag{318}$$

where R_t is the resistance of the conductor at the temperature t;
R_0 is the resistance of the conductor at zero of the temperature scale used; and
α_0, β_0, and λ_0 are coefficients characteristic of the substance.

The values of β_0 and λ_0 are usually very small, and for most purposes it is sufficiently accurate to neglect them. The relation then becomes:

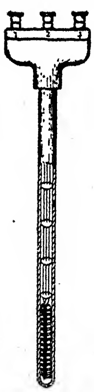


FIG. 362. Resistance Thermometer. (Courtesy of Prof. J. M. Cork)

$$R_t = R_0(1 + \alpha_0 t). \tag{319}$$

α_0 is called the **temperature coefficient of resistance.** For metals it is positive, and in most cases about 0.004 (ohm/ohm-°C). For carbon and electrolytes it is negative.

The **resistance thermometer** (Fig. 362), one of the best modern devices for temperature determinations, is made possible by these relations. It consists of a coil of fine wire, usually of platinum or nickel, whose resistance R_t is measured at any temperature that is safely below the melting point of the metal. Since R_0 and the coefficients are known for the coil from previous tests, the temperature t may be computed by Eq. (318). By this method temperature

changes of 10^{-7}C have been observed (Kimball). The platinum resistance thermometer gives a sufficiently close approximation to hydrogen scale temperatures for most purposes.

RESISTIVITIES AND TEMPERATURE COEFFICIENTS AT 0°C *

Substance	ρ_0 ohm-cm	$\alpha_0(°C)^{-1}$
Silver	1.47×10^{-6}	0.00400
Copper (annealed)	1.589×10^{-6}	0.00427
Gold	2.20×10^{-6}	0.00368
Aluminum	2.63×10^{-6}	0.00445
Nickel	6.93×10^{-6}	0.006
Platinum	11.0×10^{-6}	0.00367
Iron	11.8×10^{-6}	0.00423
Manganin	42.0×10^{-6}	0.000011
Mercury	94.07×10^{-6}	0.0008649
Hard Rubber	$1. \times 10^{18}$	

Relation of temperature coefficient to its base resistance. The resistance R_t of a conductor at the temperature t may be computed from its resistance R_1 at *any* temperature t_1 as a base by the formula:

$$R_t = R_1[1 + \alpha_1(t - t_1)] \tag{320}$$

instead of Eq. (319). It may be shown easily that the temperature coefficients α_0 and α_1 are connected with their corresponding base resistances R_0 and R_1 by the relation:

$$R_0\alpha_0 = R_1\alpha_1 = R_2\alpha_2 = \cdot\cdot\cdot. \tag{321}$$

412. Superconductivity.† The resistance of most metals such as platinum, gold, copper, and iron decreases (but not linearly) as the temperature decreases, and *gradually* approaches zero as the temperature approaches absolute zero.

But Kammerlingh Onnes found in 1911 that at 4.22°K, the resistance of mercury *suddenly* dropped to practically zero (actually to 10^{-12} of its value at 0°C). **This phenomenon of suddenly losing all resistance at very low temperatures, he called superconductivity.**

The property is exhibited by lead, tin, tantalum, niobium, in-

* Chiefly from Pender, *Handbook for Electrical Engineers.*
† E. F. Burton, *Superconductivity* (University of Toronto Press, 1933).

dium, and a few other pure metals, which are called **superconductors.** Various alloys show the phenomenon, notably one of gold and bismuth, neither of which is itself a superconductor.

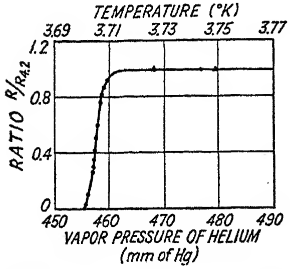


FIG. 363. Curve Showing Superconductivity

Figure 363 shows the behavior of lead with regard to resistance as the temperature approaches absolute zero.

Superconductivity has not yet been explained. To demonstrate it, a lead ring is supported so as to embrace a magnetic flux and is then submerged in liquid helium. On destroying the magnetic flux, there is induced in the ring a current which persists for days, as may be shown by the deflection of a neighboring compass needle.

413. Resistances in series. When a given number of electrons pass per second through several conductors in series, it would seem that the total resistance R that they encounter will be the sum of the resistances R_1, R_2, R_3, etc., that they encounter in the individual conductors respectively. That is,

$$R = R_1 + R_2 + R_3 + \cdots$$

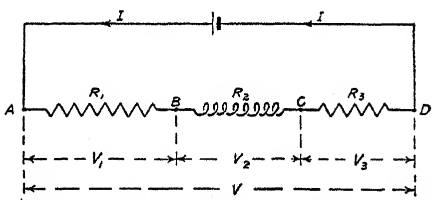


FIG. 364. Resistances in Series

This may be easily proved. Consider several resistances, R_1, R_2, R_3, etc., in series in a circuit as shown in Fig. 364.

Let R be the total resistance from A to D by the route $ABCD$;
V be the potential difference between A and D;
V_1 be the potential difference between A and B;
V_2 be the potential difference between B and C; and
V_3 be the potential difference between C and D.

Since the potential difference between two points is independent of the path by which the test charge is moved from the point of lower to the point of higher potential (Sec. 353), it follows that

$$V = V_1 + V_2 + V_3 + \cdots \qquad \text{(a)}$$

By the law of continuity of current, the current has the same value I at all points of the circuit. And by Ohm's law,

$$V = RI$$
$$V_1 = R_1 I$$
$$V_2 = R_2 I$$
$$V_3 = R_3 I, \text{ etc.}$$

Substituting these values in Eq. (a),

$$RI = R_1 I + R_2 I + R_3 I + \cdots$$

Dividing through by I,

$$R = R_1 + R_2 + R_3 + \cdots \tag{322}$$

In the figure are shown the *conventions* for representing the different kinds of resistances: R_1 and R_2 represent **noninductive** resistances; R_2 represents **inductive** resistance.

414. Resistances in parallel. When several resistances are connected in parallel, any one is said to be a **shunt** to all the others. If a given number of electrons pass per second through a system of resistances in parallel (Fig. 365), they will divide up into groups in which the numbers of electrons are in inverse proportion to the resistances of the paths. Hence the whole original number does not encounter the resistances of all the branches; only the group of electrons through a given branch will encounter the resistance of that branch.

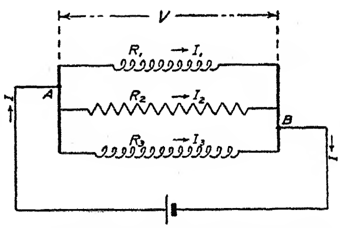


FIG. 365. Resistances in Parallel

Let R be the single **equivalent** resistance of the parallel system consisting of R_1, R_2, R_3, etc. That is, if the parallel system were entirely removed and R connected in its place between points A and B, the current I in the circuit as a whole would not be changed.

The end connectors are supposed to be so large that their resistances are negligible; hence the potential difference between the ends of R_1, R_2, R_3, etc., is in each case V, the potential difference between the end connectors.

By the law of continuity of current,

$$I = I_1 + I_2 + I_3 + \cdots \tag{a}$$

And by Ohm's law,

$$I = \frac{V}{R}, \quad I_2 = \frac{V}{R_2},$$

$$I_1 = \frac{V}{R_1}, \quad I_3 = \frac{V}{R_3}, \text{ etc.}$$

Substituting these values of the I's in Eq. (a),

$$\frac{V}{R} = \frac{V}{R_1} + \frac{V}{R_2} + \frac{V}{R_3} + \cdots$$

or

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots \tag{323}$$

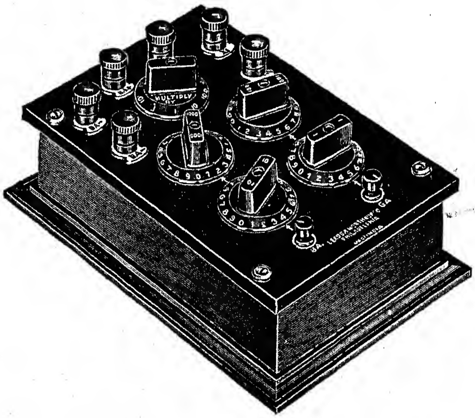


FIG. 366. Wheatstone Bridge. (Courtesy Leeds and Northrup Co.)

which is the required relation between the single equivalent resistance R and the individual resistances R_1, R_2, R_3, etc., of a parallel system.

415. The Wheatstone-Christie bridge. The most generally used device for determining resistances is commonly called the Wheatstone bridge after Sir Charles Wheatstone, who always insisted that it was devised by Hunter Christie (1843). A modern design of this bridge is shown in Fig. 366. The wiring diagram for it is shown in Fig. 367.

The unknown resistance R_4 is connected in series with three known resistances R_1, R_2, and R_3, to form a closed circuit, which is conveniently represented in the form of a parallelogram $ABCD$.

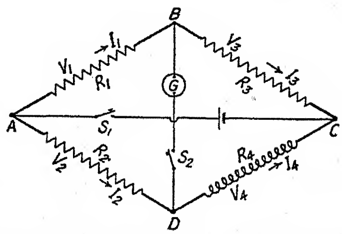


FIG. 367. Diagram of Wheatstone Bridge Connections

The resistances R_1 and R_2 are adjustable so that their ratio may be 0.001, 0.01, 0.1, 1, 10, 100, or 1000; and they are called the

"ratio arms" since only their ratios are used. R_3 is adjustable over a wide range of values in steps of 1 ohm, and is called the "adjustable arm." R_4 is the resistance to be determined. A galvanometer G is connnected in the vertical diagonal, and a cell in the horizontal diagonal, though the cell and the galvanometer are interchangeable.

R_1 and R_2 are set at convenient values, and R_3 is then adjusted until, on closing *first* S_1 and then S_2, the galvanometer shows no deflection. The bridge is then said to be "balanced," and the relation among the resistances is:

$$\frac{R_1}{R_2} = \frac{R_3}{R_4}.$$

The proof is as follows. When the galvanometer shows no deflection, there is no current in it. Hence the potential at B must equal the potential at D, for otherwise there would be a flow of electricity either downward or upward in that diagonal. Therefore,

Poten. diff. between A and B = Poten. diff. between A and D, or

$$V_1 = V_2. \tag{a}$$

Similarly,

Poten. diff. between B and C = Poten. diff. between D and C, or

$$V_3 = V_4. \tag{b}$$

But by Ohm's law,

$$\left.\begin{aligned} V_1 &= R_1 I_1 \\ V_2 &= R_2 I_2 \\ V_3 &= R_3 I_3 \\ V_4 &= R_4 I_4 \end{aligned}\right\} \tag{c}$$

Substituting from Eq. (c) in Eqs. (a) and (b);

$$R_1 I_1 = R_2 I_2 \tag{d}$$

$$R_3 I_3 = R_4 I_4. \tag{e}$$

But again, since no electricity flows through the galvanometer,

$$I_1 = I_3 \tag{f}$$

and

$$I_2 = I_4. \tag{g}$$

Hence, Eq. (e) becomes

$$R_3 I_1 = R_4 I_2. \tag{h}$$

And dividing Eq. (d) by Eq. (h),

$$\frac{R_1 I_1}{R_3 I_1} = \frac{R_2 I_2}{R_4 I_2}$$

so that

$$\frac{R_1}{R_3} = \frac{R_2}{R_4} \tag{324}$$

and by alternation,

$$\frac{R_1}{R_2} = \frac{R_3}{R_4}. \tag{325}$$

416. Fall of potential along a conductor. Consider a conductor AD in which a current I is maintained by a cell, as shown in Fig. 368.

Since electricity flows from A to B to C to D, the potential at A must be higher than at B; at B higher than at C; and at C higher than at D—because electricity flows from points of higher to points of lower potential.

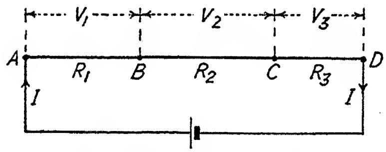


FIG. 368. Potential Decreases along a Conductor

That is, as we go along a conductor in the direction of the current, the potential decreases, or falls off, or drops off. The amount by which the potential decreases from one point to another is therefore called the **fall of potential,** the **potential drop,** or the **ri-drop** between the two points.

If the resistances between the different points are as shown in Fig. 368,

$$\left.\begin{array}{ll}\text{the fall of potential between } A \text{ and } B \text{ is} & V_1 = R_1 I \\ \text{the fall of potential between } B \text{ and } C \text{ is} & V_2 = R_2 I \\ \text{the fall of potential between } C \text{ and } D \text{ is} & V_3 = R_3 I\end{array}\right\} \text{by Ohm's law.}$$

In each instance the potential drop V has the form RI; hence the name "ri-drop," mentioned above.

By taking the ratios of the corresponding sides of the above equations,

$$V_1 : V_2 : V_3 = R_1 : R_2 : R_3 \tag{326}$$

from which it is seen that **the fall of potential from point to point along a conductor is proportional to the corresponding resistance between the points.**

417. The potentiometer. One of the most important instruments in a modern research laboratory is the potentiometer. It is a device for determining an unknown potential difference **by balancing** it against a known potential difference.

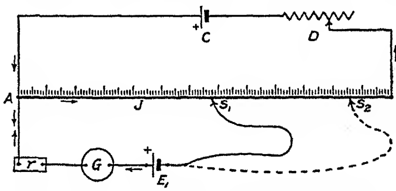


FIG. 369. Diagram of Potentiometer

The potentiometer consists of a long slide-wire AB (Fig. 369), or its equivalent, of uniform composition and cross section, often stretched beside a scale of length. A current I is maintained in the main circuit ABC by the cell C, preferably a lead storage cell. ***I* must be kept constant** by adjusting the rheostat D when necessary.

A **standard cell** of known emf E_1 is connected in a side circuit with a galvanometer G and a protective resistance r, which may be cut out before final precise adjustment.

Since electricity is flowing from A to B, any point between A and B is at lower potential than A. Hence, if the sliding contact is set at any point J, electricity will tend to flow from A to J both along the slide-wire and along any other path such as the side circuit AGE_1J.

But in the latter circuit, current from A is opposed by the emf E_1, provided the standard cell is connected with its polarity as shown. It is clearly possible to move the sliding contact along AB until a point S_1 is found, such that **in the side circuit** the tendency of the cell E_1 to produce a current to the left is exactly counterbalanced by the tendency of the potential difference between A and S_1 to produce a current to the right.

Since there will then be **no current** in the side circuit, this "balanced" condition is indicated by **zero deflection of the galvanometer,*** and the emf E_1 of the standard cell must be equal to the ri-drop from A to S_1, since they balance each other. That is,

$$E_1 = R_1 I \tag{a}$$

* Such methods are called zero deflection, or null methods.

where R_1 is the resistance of the slide-wire from A to S_1, length L_1 cm.

Let any unknown potential difference, such as the emf E_2 of another cell, be put in the place of E_1, and the sliding contact moved until this new emf E_2 is balanced by the potential difference from A to S_2, length L_2 cm.

Then,

$$E_2 = R_2 I \tag{b}$$

where R_2 is the resistance of the slide-wire from A to S_2.

Dividing Eq. (b) by Eq. (a),

$$\frac{E_2}{E_1} = \frac{R_2}{R_1}.$$

But by Eq. (316),

$$R_2 = \rho\frac{L_2}{A} \text{ and } R_1 = \rho\frac{L_1}{A}.$$

Therefore,

$$\frac{E_2}{E_1} = \frac{\rho\dfrac{L_2}{A}}{\rho\dfrac{L_1}{A}} = \frac{L_2}{L_1}$$

or,

$$E_2 = \left(\frac{E_1}{L_1}\right)L_2 \tag{327}$$

where E_1, L_1, and L_2 are either known or observed.

The factor $\left(\frac{E_1}{L_1}\right)$ is obviously the fall of potential along the slide-wire per unit of length, and will be constant if the wire is uniform in cross section and composition. Consequently, when $\left(\frac{E_1}{L_1}\right)$ has been determined, the apparatus is calibrated as a potentiometer. Then any length L_2 required to balance any unknown potential difference, when multiplied by the **calibration factor** $\left(\frac{E_1}{L_1}\right)$, gives the value of the unknown potential difference.

A commercial form of potentiometer is shown in Fig. 370. These instruments may be calibrated to read directly in volts, pH units, temperatures, or any other quantity which varies with emf.

A potentiometer measures the **total emf** of a source or the **total**

potential difference between two points; because when the measurement is made, there is no current in the galvanometer circuit, hence there is no ri-drop in the source under test or in the connecting wires.

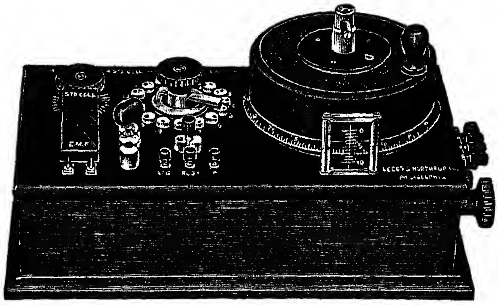


FIG. 370. Potentiometer. (Courtesy Leeds and Northrup Co.)

A voltmeter, on the other hand, requires a small current to actuate it. On account of this current there is always a corresponding ri-drop in the source of emf and in the connecting wires; and the **voltmeter reading is less than the true value** by this amount, since the voltmeter registers only the potential difference between its own terminals.

PROBLEMS

1. The resistivity of pure annealed copper (standard) at 0°C is 1.589×10^{-6} ohm-cm. Compute the resistance of a wire 0.5 mm in diameter and 10 m long.

2. If the diameter of No. 24 copper wire is 20.1 mils at 0°C, find the resistance of 100 ft at that temperature.

3. The diameter of No. 20 copper wire is 31.9 mils at 0°C. Find the resistance of 200 ft at that temperature.

4. The resistance of a mil-foot of copper wire at 20°C is 10.371 ohm. Find the resistivity of copper in ohm-cm at 20°C.

5. The resistance of a mil-foot of German silver wire is 150 ohms. What is its resistivity in ohms/cm^3?

6. If the specific resistance of aluminum is 2.828×10^{-6} at 20°C, compute the resistance of 100 ft of aluminum wire 1/20 in. in diameter.

7. If the resistance of a platinum resistance thermometer at 0°C is 12 ohms, what is its resistance at 1000°C?

8. What is the temperature of the above thermometer when its resistance is 14.5 ohms?

9. If the resistance of a nickel resistance thermometer at 0°C is 15 ohms, what is its temperature when its resistance is 18.4 ohms?

10. The resistance of a mil-foot of copper wire at 20°C is 10.371 ohms. Find its resistance at 0°C.

11. Find the resistance of a mil-foot of copper wire at 100°C.

12. What is the temperature of the platinum resistance thermometer of problem 7 when its resistance is 60.44 ohms?

13. The resistance of the copper winding of a generator at room temperature (20°C) is 0.033 ohm. What is its temperature in operation if its resistance is then 0.041 ohm?

14. The resistance of the copper winding of a generator at room temperature (15°C) is 0.042 ohm. What is its temperature in operation if its resistance is then 0.051 ohm?

15. A slide-wire is connected in circuit with a lead storage cell to form a potentiometer. A cadmium standard cell (emf 1.0183 volts) is placed in the shunt circuit with the galvanometer and is balanced by the ri-drop over 50.91 cm of the slide-wire. The standard cell is then replaced by another cell whose emf is balanced by the ri-drop over 76.63 cm. What is the ri-drop per cm along the wire, and what is the emf of the second cell?

CHAPTER XXVIII

ELECTRIC ENERGY AND POWER

418. Electric work: Weber's law. In Sec. 401, electromotive force was defined as the work required to carry a unit charge around an electric circuit. Thus far, however, we have not shown that the E in the law of Henry and Faraday conforms to this definition. It is easily done.

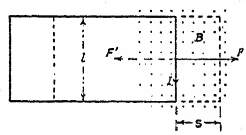


FIG. 371

In Fig. 371, let a rectangular coil of N turns be moved so that only one side l cm long cuts at right angles across a uniform magnetic field of B lines/cm², moving a distance s cm in t sec; and let mechanical friction be negligible.

According to the law of Henry and Faraday, an average emf (E abvolts) will be developed in the coil, such that

$$E = N\frac{\Phi_2 - \Phi_1}{t} = N\frac{\Phi}{t}.^* \tag{328}$$

This emf will produce a current I abamperes in the coil and in consequence it will be acted upon by a force F' dynes, whose value is given by Ampère's law:

$$F' = NBlI \qquad \text{toward the left.}$$

Hence, to keep the coil moving across the magnetic field will require an external force F dynes, at least as great as F'. That is,

$$F = NBlI \qquad \text{toward the right.}$$

The work W ergs done by the external force in moving the coil the distance s is:

$$\begin{aligned} W &= Fs \cos\theta = NBlIs \qquad (\text{since } \cos\theta = 1) \\ &= NBAI \qquad (\text{where } A \equiv ls = \text{area swept over by } l) \\ &= N\Phi I. \end{aligned} \tag{329}$$

* The $-$ is omitted here because we are interested in magnitude only.

This relation, known as *Weber's law*, states that to move a current I abamperes in a coil of N turns across Φ lines of magnetic flux requires $N\Phi I$ ergs of work.

From Eq. (328) above we have:

$$N\Phi = Et$$

which, with Eq. (329) gives:

$$W = EIt. \tag{330}$$

And since, from Eq. (295),

$$It = Q$$

$$W = EQ$$

and

$$E = \frac{W}{Q}.$$

Thus the E in the law of Henry and Faraday is work per unit charge. Since frictionless conditions were assumed, no work is wasted in friction; hence all the work done by the external force F must be utilized in maintaining the current in the coil. E must therefore be the work done on unit quantity of electricity to make it move around the circuit; and E conforms to the definition of emf, Eq. (305).

In Eq. (330), W is the energy supplied to the circuit by the generator. It may be utilized in driving a motor, heating a lamp filament, or in any other way desired. Though derived for emu, it is equally correct for practical units, as is easily shown.

Let E volts and I amp be the emf and current, respectively. Then

$$E \text{ volts} = 10^8\, E \text{ abvolts} \qquad \text{by definition}$$

and

$$I \text{ amperes} = \tfrac{1}{10} I \text{ abampere} \qquad \text{by definition.}$$

Writing these values in Eq. (330):

$$\begin{aligned}(E \text{ volts})(I \text{ amperes})(t \text{ sec}) &= (10^8 E \text{ abv})(\tfrac{1}{10} I \text{ abamp})(t \text{ sec}) \\ &= 10^7 (E \text{ abvolts})(I \text{ abamp})(t \text{ sec}) \\ &= 10^7\, W \text{ ergs} \qquad \text{by Eq. 332} \\ &= W \text{ joules.} \end{aligned} \tag{331}$$

Hence Eq. (330) is correct for either emu or practical units. This of course is **not accidental,** for the joule was defined in the first place so that this relation would be true in both systems.

Similarly, the work done by the electricity in any part of a circuit is:

$$W = VIt \tag{332}$$

where V is the fall of potential, or potential drop, across that part of the circuit; I the current in it; and t the time; and the relation is correct for both emu and practical units.

Engineers generally do not employ different symbols for emf and potential drop, but make the distinction between E and V mentally.

419. Joule's law. In case all the energy supplied to a part of a circuit, or to the whole circuit, is used in overcoming resistance, being thereby transformed into heat, it is convenient to have it expressed in terms of the resistance R in which the current is maintained.

By Ohm's law,

$$V = RI.$$

Substituting in Eq. (332),

$$W = (RI)It$$

or

$$W = RI^2t \tag{333}$$

where

W is in ergs when R is in abohms, I in abamp, and t in sec;

or

W is in joules when R is in ohms, I in amp, and t in sec.

From the first law of thermodynamics, we know that when a given amount of energy is expressed in units of work as W, and when converted into heat yields H units of heat, there is the relation:

$$W = JH \tag{334}$$

where J is known as Joule's equivalent, or *the mechanical equivalent of heat.*

When W is in joules and H in calories, the value of J is

$$J = 4.185 \text{ joules/calorie.}$$

Substituting in Eq. (333),

$$JH = RI^2t$$

$$H = \frac{1}{4.185}\, RI^2t$$

$$= 0.24\, RI^2t. \tag{335}$$

This relation was obtained experimentally by Joule in 1841, and is known as *Joule's law*.

It should be noted that Eq. (335) is correct only for practical units, on account of the constant 0.24.

420. Electric power. Power P has been defined in Sec. 82 as

$$P \equiv \frac{W}{t}$$

and from Eq. (330),

$$W = EIt.$$

Therefore,

$$P = EI \tag{336}$$

is the expression for the power developed by a generator and delivered to a circuit. Similarly, the power expended in a part of a circuit is:

$$P = VI \tag{337}$$

where V is the fall of potential across that part of the circuit.

In case all the power is employed in overcoming resistance, we express P in terms of R.

By Ohm's law,

$$V = RI.$$

Therefore

$$P = (RI)I = RI^2. \tag{338}$$

In this case all the energy degrades into heat. If that heat is wasted, as in transmission lines, **this power is called the "RI^2 loss."**

These equations are obviously correct for all emu or all practical units. If E is in abvolts, I in abamperes, and R in abohms, P is in ergs/sec. If E is in volts, I in amp, and R in ohms, P is in joules/sec, or watts.

421. Other units of work and power. Besides the erg per second, and the watt ($\equiv$ 1 joule per second), other units of power in common use are:

$$\left.\begin{aligned} 1 \text{ kilowatt} &\equiv 1000 \text{ watts} \\ 1 \text{ microwatt} &\equiv 10^{-6} \text{ watt} \end{aligned}\right\} \tag{339}$$

The **watt-hour** is defined as the **work** done when energy is used at the rate of one watt for a period of one hour.

$$\begin{aligned} 1 \text{ watt-hour} &\equiv 1 \text{ watt} \times 1 \text{ hour} \\ &= 60 \text{ watt-minutes} \\ &= 3600 \text{ watt-seconds} \\ &= 3600 \times \frac{1 \text{ joule}}{1 \cancel{\text{sec}}} \times 1 \cancel{\text{sec}} \\ &= 3600 \text{ joules} \end{aligned} \tag{340}$$

Therefore

$$1 \text{ kilowatt-hour} = 3{,}600{,}000 \text{ joules} \tag{341}$$

Solved Problems

1. A certain pressing iron has a resistance of 20 ohms. What power will it use when connected across 110-volt mains? What will it cost to operate it 2 hr per day for a month, if the price of electric energy is 5¢ per kilowatt-hour?

Known:

$$\begin{aligned} R &= 20 \text{ ohms} \\ E &= 110 \text{ volts} \\ t &= 2 \text{ hr per day} = 60 \text{ hr per month} \\ \text{Rate} &= \$.05 \text{ per kwh.} \end{aligned}$$

Solution:

$$I = \frac{E}{R} = \frac{110}{20} = 5.5 \text{ amp}$$

$$\begin{aligned} P &= EI = 110 \times 5.5 = 605 \text{ watts} \\ \text{Energy per month} &= 605 \times 60 = 36{,}300 \text{ watt-hours} \\ &= 36.3 \text{ kwh} \\ \text{Cost} &= 36.3 \times .05 = \$1.82. \end{aligned}$$

2. A motor whose input power is 1 kw operates at 110 volts. If the loss in the line is 3%, what is the current, the voltage at the generator, and the resistance of the line?

Known:

$$\begin{aligned} \text{Power at motor, } P_2 &= 1 \text{ kw} = 1000 \text{ watts} \\ \text{Voltage at motor, } E_2 &= 110 \text{ volts} \\ \text{Power lost in line} &= 3\%. \end{aligned}$$

Solution:

$$P_2 = E_2 I$$
$$1000 = 110 \times I$$
$$I = \frac{1000}{110} = 9.09 \text{ amp}$$
$$P_2 = .97\ P_1, \text{ the power delivered by generator}$$
$$P_1 = \frac{1000}{.97} = 1030.93 \text{ watts} = \text{power developed by generator}$$
$$P_1 = E_1 I$$
$$E_1 = \frac{1030.93}{9.09} = 113.4 \text{ volts.}$$
$$\text{Power lost in line} = RI^2 = 0.03 \times 1030.9 = 30.927 \text{ watts}$$
$$R = \frac{30.93}{(9.09)^2} = 0.374 \text{ ohm.}$$

PROBLEMS

1. When there is a current of 4 amp in a resistance of 12 ohms, how much heat in joules per minute is developed?

2. When a current of 5 amp is in a soldering iron whose resistance is 15 ohms, how much heat in calories is developed per minute?

3. It is required to generate 2000 calories of heat per minute in a circuit. The potential difference of the terminals of the circuit is 200 volts. What resistance must be used?

4. It is required to generate 1500 calories of heat per minute in a circuit. The potential difference of the terminals of the circuit is 100 volts. What resistance must be used?

5. How many minutes are required to heat 800 gm of water from 15° to 100°C by passing 8 amp through a 12-ohm coil immersed in water, if 20% is lost by radiation?

6. What current will a 0.8-kw heating unit of an electric range take from 110-volt mains?

7. How many minutes are required to heat 1000 gm of water from 20° to 100°C by passing 10 amp through a 10-ohm coil immersed in water, if 15% is lost by radiation?

8. What current will a 0.6-kw heating unit of an electric range take from 22-volt mains?

9. When a current of 4 amp is in a soldering iron whose resistance is 20 ohms, how much heat in joules is developed per minute? In calories?

10. The ordinary 16-candle-power tungsten lamp has a resistance when hot of 440 ohms. How much power is required to operate 4 of these in parallel across 110-volt mains?

11. If the above lamps are connected in series and energized so as to produce the same amount of light as before (i.e., have the same current maintained in them), what voltage and what power would be required?

12. What current will a 0.5-kw heating unit of an electric range take from 110-volt mains?

13. A generator develops 50 amp at 100 volts. What is the power in kilowatts and what in horsepower?

14. A 10-hp engine is to be replaced by an electric motor. What must be the power of the motor? If it is to be connected to 220-volt mains and its efficiency is 90%, what current will it take?

15. An electric motor requires 10 amp at 100 volts. The two wires leading from the dynamo have a resistance of 0.3 ohm each. What emf must the dynamo supply? How much power is lost in the line?

16. A generator whose terminal voltage is 220 volts produces a current of 6 amp in the power line to a motor. If the resistance of each wire is 0.2 ohm, how much power is delivered to the motor? How much is wasted as heat?

CHAPTER XXIX

ELECTROLYSIS—BATTERIES

422. Electrolysis is the process of decomposing a substance by passing electricity through it. It was discovered in 1799 by the German chemist, J. W. Ritter, in connection with a drop of water. The following year he showed that the gases evolved are hydrogen and oxygen.

Not all solutions can be electrolyzed, but certain substances in the fused condition can be. Thus, metallic potassium and sodium were first obtained by Sir Humphry Davy in 1807 by electrolysis of their molten hydroxides at about 380°C. Fluorine is obtained by electrolysis of potassium fluoride maintained in the liquid state in an electric furnace.

Faraday, who made the most searching study of the phenomenon, gave it the name *electrolysis*. To him also are due the names:

Electrolyte—an electrolytic conductor, usually a solution which is ionized and which is decomposed by the passage of electricity through it.

Anode (Greek, the way in)—the electrode, or conductor, by which electricity enters the electrolyte.

Cathode (Greek, the way out)—the electrode, or conductor, by which electricity leaves the electrolyte.

Ion (Greek, goer)—an atom or group of atoms carrying a charge of electricity and which, in consequence, may move through an electrolyte.

423. Electrolytic conduction. Electrolytic conduction is quite different from metallic conduction (Sec. 409). When various substances are dissolved in certain solvents, the forces holding together the groups of atoms that make up the molecules are weakened to such an extent that these groups of atoms with their associated charges are able to *go* around in the solution quite freely, and are hence called **ions** (goers). A substance broken up into such groups is said to be **ionized.**

Many molten substances conduct electrolytically, and even some solid salts such as silver iodide, but the conduction of liquid mercury is metallic. Substances which conduct electrolytically are decomposed in the process.

The conductivity depends both upon the solvent and the solute. **Solutions of acids, bases, and salts in water are electrolytes,** i.e., are ionized; but a water solution of sugar is not. Some solutions in alcohol are fairly good conductors. But generally speaking, solutions in organic solvents are poor conductors, whereas those in liquid SO_2, NH_3, and HCN are good.

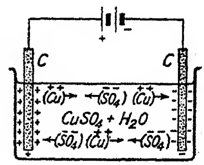


FIG. 372. Electrolytic Conduction

In a water solution of $CuSO_4$, for example, the salt is broken up into ions of SO_4^{--} carrying a charge of two electrons, and ions of Cu^{++} carrying two positive charges each equal in magnitude to an electron (Fig. 372).

If inert electrodes, such as carbon or platinum, are connected to a battery, electrons will be withdrawn from one, making that one positive, and given to the other, which is therefore negative. When these electrodes are dipped into the solution, the Cu^{++} ions will be attracted to the negatively charged electrode (cathode), where their two + charges will be neutralized by two − charges from the electrode; and metallic copper will be deposited on the electrode. Similarly, the SO_4^{--} ion will be attracted to the positively charged electrode (anode), where its two electrons will be neutralized by two positive charges from that electrode. The SO_4 radical, being then unstable, will break down the solvent H_2O, liberating O_2 in bubbles at the anode and forming H_2SO_4 with the H_2. Copper is therefore removed from the solution, which gradually becomes H_2SO_4. If the anode were of copper, the SO_4 would rob it of an atom of Cu to form $CuSO_4$, and the concentration of the solution would remain constant.

It will be seen that in electrolytic conduction the electricity is transferred by two streams of oppositely charged ions going through the solution in opposite directions, whereas in metallic conduction there is only one stream of electrons moving in one direction.

The positive ion is called the **cation** because it goes to the cathode; and the negative ion, the **anion** because it goes to the

anode. The **hydrogen ion,** like metallic ions, always carries a + charge and moves through the solution **in the direction of positive current.** Hence these elements are always liberated at the cathode.

424. Faraday's laws of electrolysis. After his epochal discovery of electromagnetic induction, Faraday took up the study of conduction of electricity through electrolytes and in 1834 evolved the following laws:

1. The mass M of a substance liberated by electrolysis is proportional to the quantity Q of electricity that passes through the electrolyte.

2. When the same quantity Q of electricity passes through different electrolytes, the masses of the various substances liberated are proportional to their chemical equivalents, or combining masses. The chemical equivalent, or combining mass, of an element is defined as its atomic mass divided by its valence.

The first law expressed algebraically is:

$$M \propto Q.$$

Therefore

$$M = zQ \tag{342}$$

and since, by Eq. (295), $Q = It$

$$M = zIt \tag{343}$$

where the constant of proportionality z depends upon the units used and the substance liberated. The constant z is called the **electrochemical equivalent** of the substance liberated, and is a characteristic property of that substance.

Solving Eq. (343) for z,

$$z \equiv \frac{M}{It} \equiv \frac{M}{Q}. \tag{344}$$

Hence, when M is in grams and Q in coulombs, z is in grams per coulomb.

Stated in words, **the electrochemical equivalent of a substance is the mass in grams of that substance liberated by the passage of one coulomb of electricity through an electrolyte.**

Care should be taken to distinguish between the terms "chemical equivalent" and "electrochemical equivalent." Calling the

atomic mass m and the valence v, the chemical equivalent as defined above is m/v. From Eq. 347 it will be seen that the electrochemical equivalent is equal to chemical equivalent divided by the faraday.

ELECTROCHEMICAL EQUIVALENTS *
(Grams per Coulomb)

Element	Atomic Mass	Valence	Electrochemical Equivalent
Copper	63.57	1	0.6588×10^{-3}
Copper	63.57	2	0.3294
Chromium	52.01	3	0.1796
Gold	197.2	3	0.6812
Hydrogen	1.0078	1	0.0105
Nickel	58.69	2	0.3040
Silver	107.88	1	1.1180
Zinc	65.38	2	0.3387

The most used electrochemical equivalent is perhaps that of silver, which from experiment has the value:

$$z_{Ag} = .001118 \quad \text{gm/coulomb.}$$

By combining the two laws of electrolysis algebraically, we obtain an expression which may be considered *the general law of electrolysis.*

From the first law,

$$M \propto It \qquad \text{when atomic mass } m \text{ and valence } v \text{ are constant,}$$

and from the second law,

$$M \propto \frac{m}{v} \qquad \text{when } I \text{ and } t \text{ are constant.}$$

Consequently, both laws together are expressed by the relation:

$$M \propto \left(\frac{m}{v}\right)It \qquad \text{when all vary.}$$

From this we have the *general law:*

$$M = \frac{1}{F}\left(\frac{m}{v}\right)It. \tag{345}$$

* Chemical Rubber Co., *Hand Book of Physics and Chemistry.*

The constant F is called the "Faraday constant," or the **faraday,** and is a universal constant, its value depending only upon the choice of units for the other factors. It is written in the denominator in order that its meaning may be more easily discerned. That meaning at once becomes clear on solving Eq. (345) for F:

$$F = \frac{\left(\frac{m}{v}\right)}{M} Q. \tag{346}$$

A mass in grams numerically equal to the chemical equivalent $\frac{m}{v}$ of a substance has been defined as the **gram equivalent** of that substance.

If, therefore, in a given case the mass M of a substance liberated is 1 gram equivalent $\left(= \frac{m}{v} \text{ grams}\right)$ of that substance, we have

$$M = \frac{m}{v}$$

and therefore, by Eq. (346),

$$F = Q.$$

Hence, **the faraday is the quantity of electricity in coulombs necessary to deposit one gram equivalent of any substance by electrolysis.**

To find the numerical value of F we proceed as follows.

From Eq. (343),

$$\frac{M}{It} = z$$

and from Eq. (345),

$$\frac{M}{It} = \frac{1}{F}\frac{m}{v}.$$

Therefore

$$z = \frac{1}{F}\frac{m}{v} \tag{347}$$

and

$$F = \frac{1}{z}\frac{m}{v}. \tag{348}$$

Using the data for silver,

$$m = 107.88 \text{ gm}$$
$$v = 1$$
$$z = 0.001118 \text{ gm/coulomb}$$

whence

$$F = \frac{1}{0.001118} \times \frac{107.88}{1} = 96{,}494 \text{ coulombs.}$$

Since electrochemical equivalents and atomic masses have not all been determined with the same accuracy, slightly different values of F will be found when different substances are used.

425. Discussion of Faraday's laws of electrolysis. The essence of these laws is that **each ion in an electrolyte carries a number of elementary charges equal to its valence.** We shall prove presently that the magnitude of each elementary charge is that of one electron. From this simple statement, we may deduce Faraday's laws in a logical manner.

Thus, suppose that each ion of a univalent element consists of one atom and one elementary charge, as Ag^+, H^+, Cl^-, etc. Then each elementary charge reaching an electrode would bring with it an atom of the element in question; hence the mass liberated would be proportional to the quantity of electricity passing through the electrolyte.

Likewise, if a divalent ion consists of a divalent atom and two elementary charges, as Cu^{++}, Zn^{++}, O^{--}, etc., then each two elementary charges reaching an electrode would bring with them one atom of the element in question; and for each two elementary charges passing through the electrolyte, an atom would be liberated at an electrode. Here again the mass liberated would be proportional to the quantity of electricity passing through the electrolyte, which is the first law. A similar relation may be worked out for atoms of any valence.

There are 6.064×10^{23} atoms in a gram-atom.* Hence, if we pass 6.064×10^{23} elementary + charges through a series of cells containing electrolytes of, say, $AgNO_3$, $CuSO_4$, and $Cr(NO_3)_3$, we should liberate 1 gm-atom of Ag, 1/2 gm-atom of Cu, and

* A gram-atom is defined as a mass of an element in grams numerically equal to the atomic mass of the element.

1/3 gm-atom of Cr. The masses liberated would be proportional to the atomic masses divided by the valences of the respective elements, i.e., proportional to their chemical equivalents—and this is just what is stated in Faraday's second law.

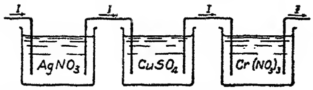


FIG. 373. Electrolytic Cells in Series

The experiment is easily made as shown in Fig. 373. Since the three cells are connected in series, the same quantity of electricity flows through each; and as predicted by the laws, the masses liberated are found to be:

$$\text{Ag} \cdots\cdots \frac{107.88}{1} = 107.88 \text{ gm}$$

$$\text{Cu} \cdots\cdots \frac{63.57}{2} = 31.79 \text{ gm}$$

$$\text{Cr} \cdots\cdots \frac{52.01}{3} = 17.34 \text{ gm.}$$

426. Battery cells. When two electrodes of different conducting materials are dipped into an electrolyte, an electromotive force is developed in the system. The system is then called a *battery cell.* The action of a battery cell may be explained as follows.

Whenever an electrode is placed in an electrolyte, there is in general a tendency for its molecules to go into solution. The measure of this tendency has been given the name **solution pressure** by Nernst. At the same time there is a tendency for its molecules already in the solution to pass out of the solution. This latter tendency is called their **osmotic pressure,** and in dilute solutions it is proportional to the concentration of the ions in question, as was shown by Van't Hoff. An electromotive force depending upon the difference between the solution pressure and the osmotic pressure is developed at the surface between the electrode and the electrolyte.

In a dilute solution of sulphuric acid in water, many of the molecules of the acid H_2SO_4 are broken up into positive ions ($2H^+$) and negative ions (SO_4^{--}). This is believed to be due to the fact that the dielectric constant of water is large (81), and on this account the force of attraction between the + and − charges within the molecule is greatly weakened, in accord with Coulomb's law.

If an electrode of pure zinc is dipped into this solution, zinc molecules, each carrying two elementary + charges, will pass into the solution on account of solution pressure, leaving behind on the zinc electrode two − charges (electrons). This will continue for a short time until osmotic pressure, the repulsion of the Zn^{++} in the solution, and the attraction of the electrons left on the zinc plate prevent additional + ions from coming into the solution, and a state of equilibrium is established. Similar action takes place when a copper electrode is dipped into the solution, but it is so slight that we may disregard it here. Although the copper is also negative with respect to the solution, its potential is higher than that of the zinc (Fig. 374).

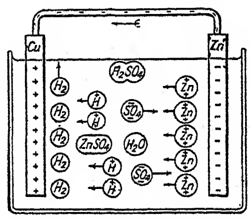


FIG. 374. Battery Cell

If the external terminals of the zinc and copper are connected by a conductor, electrons will flow from the zinc to the copper. The state of equilibrium is therefore upset and more Zn^{++} ions are released into the solution. These + ions repel the $2H^+$ ions to the copper, where their + charges are neutralized by the electrons that have come from the zinc through the external conductor, and bubbles of hydrogen gas H_2 collect on the copper and rise to the surface. Zn^{++} ions also are repelled toward the copper electrode but combine in transit with the SO_4^{--} ions to form $ZnSO_4$. This cell action continues until the zinc electrode is completely consumed.

The flow of electrons from the zinc to the copper in the external circuit is equivalent to a flow of positive charges from the copper to the zinc. As a result, the **conventional** statement that positive electricity flows from the copper to the zinc in the external circuit continues in use, although we know that actually the electrons flow the opposite way. Within the cell, positive charges travel on the ions from the zinc to the copper; hence the zinc is the **anode** and the copper, the **cathode.**

For the external circuit, the copper is called the **positive pole,** or terminal; and the zinc, the **negative pole,** or terminal.

The energy released by such a cell results from the chemical change of the zinc, just as the energy of a locomotive results from the combustion of coal. Zinc may therefore be looked upon as

the "fuel" used in this little power plant. It is obviously too expensive for large-scale power development.

Solved Problems

1. Compute the electrochemical equivalent of zinc. (Atomic mass = 65.38; valence = 2.)

Known:

$$z = \frac{1}{F}\frac{m}{v} \qquad \text{from Eq. (347)}$$
$$m = 65.38 \text{ gm}$$
$$v = 2$$
$$F = 96{,}494 \text{ coulombs.}$$

Solution:

$$z = \frac{1}{96{,}494} \times \frac{65.38}{2}$$
$$= 0.0003388 \text{ gram/coulomb.}$$

2. How long must a current of 2 amp be maintained to deposit 0.25 gm of chromium from a solution of chromic chloride, $CrCl_3$?

Known:

$$I = 2 \text{ amp}$$
$$M = 0.25 \text{ gm}$$
$$z = 0.0001796 \text{ gram/coulomb} \qquad \text{from table.}$$

Required: $t = ?$

Solution: $M = zIt.$ by Eq. (344)

Substituting values,

$$0.25 = 0.0001796 \times 2 \times t$$
$$t = 716 \text{ sec}$$
$$= 11 \text{ min } 56 \text{ sec.}$$

3. Two electrolytic cells containing $CuSO_4$ and $CdSO_4$, respectively, are connected in series in a circuit in which a current is maintained until 0.009882 gm of copper is deposited in the first cell.

How much cadmium is deposited in the second cell, if the atomic mass of cadmium is 112.41 and its valence is 2?

Known:

$$M_{Cu} = 0.009882$$
$$\text{Atomic mass } Cu = 63.57 \text{ from table}$$
$$\text{Valence } Cu = 2 \text{ from table}$$
$$\text{Atomic mass } Cd = 112.41$$
$$\text{Valence } Cd = 2.$$

Solution: Since the cells are in series, there is the same current in both. Hence the masses of Cd and Cu deposited are proportional to their chemical equivalents, by Faraday's second law of electrolysis.

Therefore,

$$\frac{M_{Cd}}{M_{Cu}} = \frac{\frac{112.41}{2}}{\frac{63.57}{2}} = \frac{112.41}{63.57}$$

$$M_{Cd} = \frac{112.41}{63.57} \times 0.009882$$

$$= 0.0175 \text{ gm.}$$

427. Defects of simple cells. There are two important defects of the simple Voltaic cell.

1. The collection of hydrogen bubbles on the positive terminal produces an effect called **polarization,** for it makes the plate behave like hydrogen rather than like copper, and this reduces the emf of the cell. By covering up the copper, this hydrogen also increases the internal resistance of the cell.

Polarization may be minimized by surrounding the positive pole with an oxidizing agent—usually manganese dioxide (MnO_2) or potassium dichromate ($K_2Cr_2O_7$)—called a **depolarizer,** which combines with the offending hydrogen.

2. **Local action** results at impurities in the electrodes which serve as one terminal of a little local cell. Thus a particle of carbon or iron in the zinc will act as the positive pole of a tiny local cell on the surface of the zinc. Electricity will flow in a short local circuit from the carbon through the zinc, and back through the electrolyte to the carbon.

Local action results in wasting of the electrode, usually the zinc, without producing any flow of electricity in the external circuit. It may be overcome to a large extent by amalgamating the zinc with mercury, which covers the impurities and dissolves the zinc and brings it to the surface.

428. Electromotive series. Since, in the simple cell, the zinc passes into solution more readily than does the copper, and each atom leaves behind on the zinc plate two electrons, we should expect the zinc plate to be the negative terminal. The experiment of Sec. 390 has already shown this to be true.

We should wonder also whether, in general, the electrode that reacts more vigorously with the electrolyte is always the negative of the cell. In order to answer this question, the potentials of

the various metals—each in a normal solution of its own ions, *taking that of hydrogen arbitrarily as zero* (see Sec. 437)—have been determined. Arranged in the order of increasing potential, they form the electromotive series, shown in the table below.

ELECTROMOTIVE SERIES

More Active		Sb	+0.1 (?)
Li	−3.02 volts	Bi	+0.2 (?)
K	−2.92	As	+0.3 (?)
Na	−2.72	Cu(++)	+0.34
Mg	−1.55	Cu(+)	+0.51
Al	−1.34	Hg(+)	+0.79
Zn	−0.76	Ag	+0.80
Fe(++)	−0.43	Hg(++)	+0.86
Ni(++)	−0.22	Pt	+0.863
Sn(++)	−0.14	Au(+)	+1.5 (?)
Pb(++)	−0.13	Less Active	
Fe(+++)	−0.04		
H_2	0.00		

By comparing this table with the chemical "activity series," * it will be seen that the order of the elements is the same in both. Hence, of two electrodes, that which is chemically more active will be the negative and the less active one, the positive electrode of a battery cell.

The electromotive force of a cell may be determined from the table by taking the difference of the electrode potentials in the table for the two metals to be used with a proper electrolyte. Thus, for the Daniell cell (see Sec. 430), the electromotive force will be that of Cu($= 0.34$) minus that of Zn($= -0.76$): $0.34 - (-0.76) = 1.10$ volts, which is very close to the usual emf of this cell.

429. Charge on a univalent ion in electrolysis: Avogadro's number. From Faraday's laws it is clear that in electrolytic conduction each ion carries a charge proportional to its valence. Thus Ag carries one elementary + charge, and Cu, two elementary ++ charges. Helmholtz suggested that this elementary quantity of electricity carried by an atom of matter might be the atom of

* See Holmes, *General Chemistry* (New York, The Macmillan Company, 1939), p. 83.

electricity, or the natural unit from which all larger quantities are made up.

Millikan's oil-drop experiment (Sec. 373) has shown convincingly that the charge of one electron is the smallest quantity of electricity that occurs in nature. If Helmholtz' hypothesis is correct, and we believe it is, this gives a method of computing **the number of atoms in a gram-atom,** or molecules in a gram-molecule, which is **Avogadro's number** N.

By Eq. (284),

$$1 \text{ electron} = 4.8025 \times 10^{-10} \text{ statcoulomb *}$$

and by experiment,

$$1 \text{ statcoulomb} = \frac{1}{2.9978 \times 10^{9}} \text{ coulomb.}$$

Hence.

$$\begin{aligned} 1 \text{ electron} &= 4.8025 \times 10^{-10} \times \frac{1}{2.9978 \times 10^{9}} \text{ coulomb} \\ &= 1.6020 \times 10^{-19} \text{ coulomb.} \end{aligned}$$

We know that one gram-equivalent, which in the case of a univalent element is also one gram-atom, is liberated by one faraday (= 96,488 coulombs) of electricity. Since each atom carries one electron, the number of atoms in 1 gram-atom, Avogadro's number, is:

$$\begin{aligned} N &= \frac{96{,}488}{1.6020 \times 10^{-19}} \frac{\text{atoms}}{\text{gram-atom}} \\ &= 6.023 \times 10^{23} \quad \frac{\text{atoms}}{\text{gram-atom}}. \end{aligned}$$

This value is in excellent agreement with that obtained by other methods, e.g., by Brownian movements.† But as the value of the electron has been determined with great accuracy, this is believed at present to be the best method of determining Avogadro's number. It confirms the hypothesis that the **magnitude of**

* Raymond T. Birge, *Rev. Mod. Phys.*, Vol. 13, No. 4, October, 1942.

† Chapin, *Second Year College Chemistry* (New York, John Wiley and Sons, 1926), p. 4.

the elementary charge on a univalent ion in electrolysis is that of one electron.

430. Primary cells. Battery cells in which the electric energy is obtained **originally** from the chemical reactions in the cell are called **primary cells.** Until about 1890, they were used to a very large extent as sources of direct current. In 1800, Sir Humphry Davy used a battery of 2000 such cells to produce the first electric arc light.

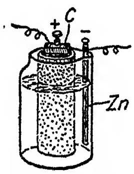


Fig. 375. Leclanché Cell

Of the many different types of primary cell that were developed, only five are now used to any great extent: the Leclanché, the dry, the Daniell, the gravity, and the Weston (cadmium standard).

1. *The Leclanché cell* (Fig. 375) has a positive pole of carbon; the electrolyte is a water solution of ammonium chloride (NH_4Cl); and the negative pole is zinc. The general reaction that takes place is as follows:

$$C + 2MnO_2 + 3NH_4Cl + 2Zn \rightarrow \overset{++}{C} + Mn_2O_3 + H_2O + 2NH_3 + NH_4Cl + ZnCl_2 + \overset{--}{Zn}.$$

Polarization, due to the liberation of H_2 at the positive pole, is diminished by surrounding this electrode with a porous jar containing manganese dioxide (MnO_2) as a depolarizer. This combines with the hydrogen to form H_2O and Mn_2O_3. However, the action is not sufficiently rapid to dispose of all the hydrogen, and the emf falls off quickly when the cell is being used continuously for several minutes. When the circuit is open, the oxidation of H_2 continues; but all other chemical activity ceases, so that the materials are not wasted.

The cell is therefore suitable for **open circuit** work, as for electric bells, which require current for but a few seconds at a time. Its emf at first is 1.5 volts; and its internal resistance is from 1 to 5 ohms, depending upon the size and condition of the cell.

2. *Dry cells* (Fig. 376) are not really dry, but they are "nonspillable." When they actually become dry they cease to function. Sometimes they may be restored to usefulness by the mere addition of water.

The dry cell is the same as the Leclanché in the chemical nature of its electrodes, electrolyte, depolarizer, and their reactions.

The zinc (negative pole) is made into a can and serves as a container. It is lined with blotting paper to prevent particles of carbon in the electrolyte from coming into contact with the zinc. On the outside it is covered with pasteboard, which insulates cells from one another.

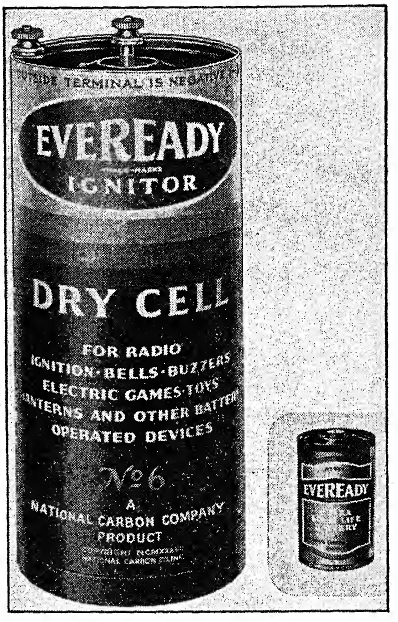


FIG. 376. Dry Cells of Equal Emf. (Courtesy National Carbon Co.)

The electrolyte (NH_4Cl) is made into a paste with flour and plaster of Paris. Powdered carbon and sawdust are added to reduce the internal resistance; MnO_2, to serve as a depolarizer; and a small amount of $ZnCl_2$, to keep the cell from becoming too dry, to prevent local action, and to combine with the NH_3 which is formed.

The positive pole is a large rod of carbon placed in the center of the cell. The top of the cell is sealed over with pitch to prevent evaporation.

The emf, which at first is 1.5 volts, decreases rapidly when a current of several amperes is produced, but recovers on resting. The internal resistance varies from 0.1 to 0.5 ohm, depending upon the size and the condition of the cell.

The emf decreases and the internal resistance increases with use. These cells become worthless from local action when they remain for a long time stored on shelves, even though not used. The time that a dry cell may stand on open circuit without its short-circuit current's falling below 10 amp is known as its "shelf life." The shelf life of good cells is from 10 to 12 months.

Millions of dry cells are now used in flashlights. Large and small, their emf is 1.5 volts at first. Greater voltages, such as for

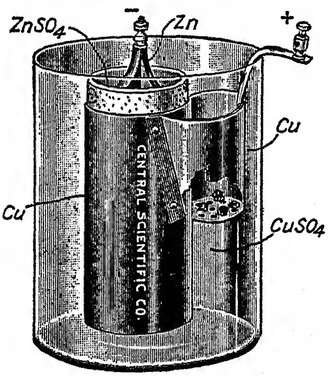


FIG. 377. Daniell Cell. (Courtesy Central Scientific Co.)

B-batteries in radio sets, are secured by connecting the proper number of these cells in series.

3. *The Daniell cell* (Fig. 377) consists of a positive pole of Cu in a saturated solution of $CuSO_4$, and a negative pole of amalgamated zinc in a 5% solution of $ZnSO_4$. The two electrolytes are separated by a porous pot of unglazed earthenware.

The reaction is substantially as follows:

$$Cu + 2CuSO_4 \Big| + ZnSO_4 + 2Zn \rightarrow 2\overset{++}{Cu} + CuSO_4 \Big| + 2ZnSO_4 + \overset{--}{Zn}.$$

This cell does not polarize, since the Cu liberated at the cathode is the same substance as the cathode itself. When on open circuit, the two electrolytes diffuse through the porous pot, wasting the $CuSO_4$ and the Zn.

The cell is therefore most suitable for **closed circuit** work. When not in use it should be kept in circuit with a resistance of 100 ohms or more, as a small current tends to keep the electrolytes from mixing.

The emf of the Daniell cell, when freshly made up, is 1.08 volts, and the internal resistance is about 0.3 to 15 ohms, depending upon its size and design.

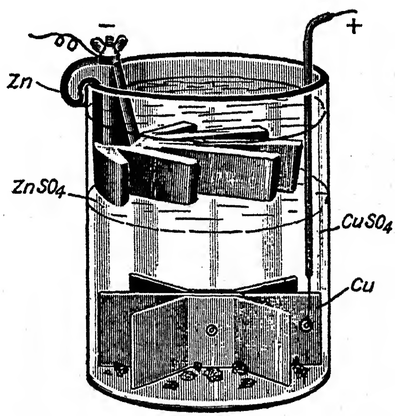


FIG. 378. Gravity Cell. (Courtesy Welch Scientific Co.)

4. *The gravity cell* (Fig. 378) is another form of Daniell cell in which the electrolytes are kept separated by gravity, the $CuSO_4$ being denser than the $ZnSO_4$. The separation is quite satisfactory so long as the cell is producing a current.

5. *The cadmium standard cell,* perfected by Edward Weston in 1893, is never used as a source of current, but only as a standard of emf. The positive electrode is of pure Hg. On top of this is a

layer of mercurous sulphate (Hg_2SO_4), which is the depolarizer. The negative electrode is an amalgam of mercury and cadmium (90Hg + 10Cd). These constituents are held in place by inert packing and porous porcelain retainers. Connections to the electrodes are made by platinum wires sealed through the glass. There are two types:

The **normal,** or **saturated,** type has as electrolyte an aqueous solution of $CdSO_4$, which is kept **saturated at all temperatures** by an excess of $CdSO_4$ crystals.

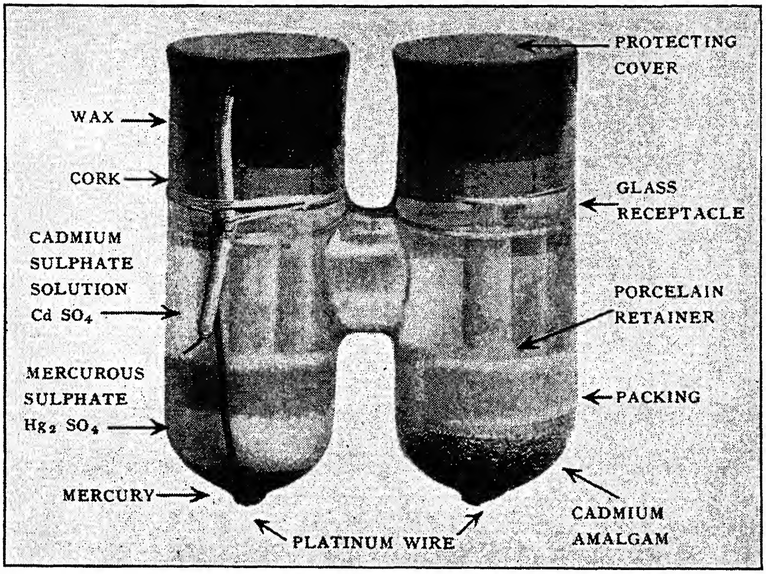


FIG. 379. Weston Cadmium Cell—Interior. (Courtesy Weston Electrical Instrument Co.)

The **unsaturated** type has an electrolyte consisting of a water solution of $CdSO_4$ saturated at 4°C. At all other temperatures it is unsaturated.

The container is usually of glass in the form of the letter H. It is hermetically sealed, either with the glass itself or with wax.

The cell proper is shown in Fig. 379. Being light-sensitive, the cell is enclosed in a light-proof cover, one form of which is shown in Fig. 380.

The normal type is the more reproducible and has been adopted by **international agreement** as the standard source of electromo-

tive force. Its emf is 1.01830 *volts at* 20°C. At other temperatures its emf is given approximately by the relation: *

$$E_t = 1.0183 - 0.00004(t - 20).$$

The unsaturated type is not so reproducible as is the normal, the emfs of different cells being from 1.0178 to 1.0192 volts. But its temperature coefficient is lower than that of the saturated type, so that within the range of temperature ordinarily encountered no temperature correction is necessary. This type is therefore used as the standard for most commercial purposes.

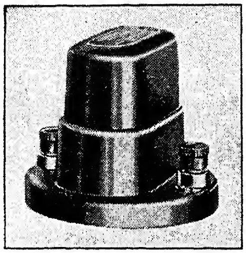


FIG. 380. Weston Cadmium Cell—Exterior. (Courtesy Weston Electrical Instrument Co.)

Cautions. A standard cell should never be permitted to transmit a current greater than 1/10,000 ampere. To provide against this, a resistance of at least 10,000 ohms should be used in series with the cell until the final adjustment in balancing.

A standard cell should not be subjected to large variations of temperature nor to rough handling.

431. Secondary, or storage, cells. Cells which are formed by the electrolytic action of the current from some other generator, and which, when discharged, can be recharged by again receiving electric energy from the generator, are called secondary cells, storage cells, or accumulators. The two principal types are: the lead-acid cell (Planté), and the nickel-iron-alkaline cell (Edison).

1. *The lead-acid cell.* Commercial cells of this type consist of a group of parallel plates of pure spongy lead (Pb), between which are interposed a group of grids of lead alloyed with a small percentage of antimony to give them strength. In the grid openings are porous packets of lead peroxide. The two sets of plates are kept apart by rubber or wooden separators, and the whole assemblage is submerged in an electrolyte of dilute H_2SO_4 in a glass or other acid-resisting container.

The brown PbO_2 plates form the positive pole; the gray Pb plates, the negative pole.

During discharge both sets of plates are partially changed to

* For the complete relation see Glazebrook, *Dictionary of Applied Physics* (London, The Macmillan Co., 1922), II, p. 273.

lead sulphate ($PbSO_4$), and the electrolyte to H_2O; on recharge the cell returns to its original condition. These reactions are represented in a general way by the following equation:

$$\begin{array}{ccc} \text{Charged} & & \text{Discharged} \\ \overset{+}{PbO_2} + 2H_2SO_4 + \overset{-}{Pb} & \rightleftarrows & PbSO_4 + 2H_2O + PbSO_4. \end{array}$$

The reactions are reversible; and provided the cell is not discharged to the extent of injuring the plates, it may be recharged repeatedly by connecting the positive terminal of the generator to the positive plates of the cell.

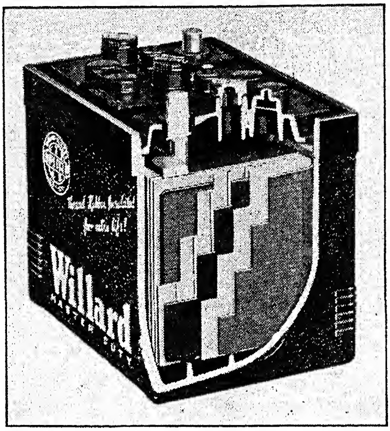


Fig. 381. Lead-Acid Storage Battery. (Courtesy Willard Storage Battery Co.)

It will be seen that, on charge, water is converted into H_2SO_4. Since the specific gravity of H_2SO_4 is greater than that of water, the specific gravity of the electrolyte as determined with a hydrometer is a convenient index of the extent to which the battery is charged.

When fully charged, the emf of the cell is 2.1 volts and the specific gravity of the electrolyte about 1.285. By the use of very thin plates very close together, the internal resistance may be made extremely small. On account of their high emf and low internal resistance, lead-acid storage cells are generally used for starting automobiles. The internal resistance of the ordinary 3-cell automobile battery (Fig. 381) is about 0.003 ohm, but larger batteries for starting Diesel engines have a resistance of about 0.001 ohm at 80°F. At this temperature the current required by the starting motor of an automobile is approximately 150 amp. At 10°F, the current required is about 300 amp.

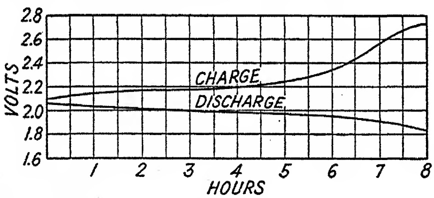


Fig. 382. Charge and Discharge Curves of Lead-Acid Cell

The energy efficiency of a lead-acid cell is from 75 to 85%.

From the discharge curve of Fig. 382, it will be observed that after discharge has continued at the normal rate for about two hours, the curve becomes nearly horizontal. At this stage of discharge, if the current is small, the emf changes very little during a considerable time; consequently, this cell is then particularly good as the main battery of a potentiometer.

The lead-acid cell has the disadvantage, however, of not being mechanically strong. It must therefore be used with these precautions:

1. Do not short-circuit.
2. Do not discharge below 1.8 volts or specific gravity 1.15.
3. Do not overcharge.
4. Do not allow the electrolyte to get below top of plates.

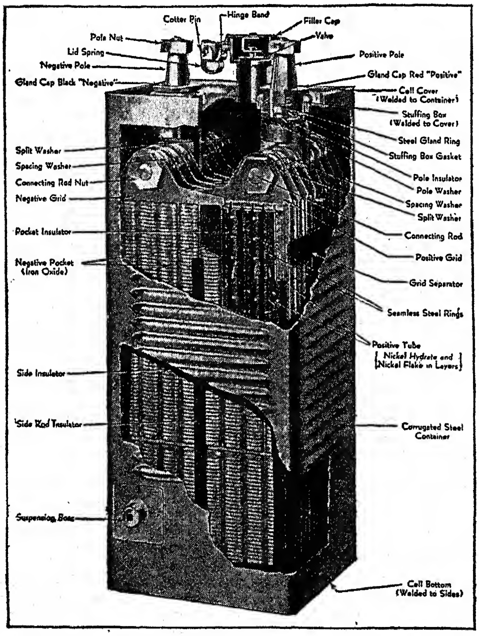


Fig. 383. Edison Storage Cell. (Courtesy Edison Storage Battery Corp.)

2. *The nickel-iron-alkaline cell*, developed in the laboratory of Thomas A. Edison, is commonly called the Edison storage cell (Fig. 383).

The electrolyte is a 21% solution of KOH in distilled water, with a small percentage of LiOH.

The positive plates consist of steel grids holding perforated cylindrical steel tubes which contain alternate layers of NiO_2 (the active material) and nickel flakes to increase conductivity.

The negative plates are a mixture of finely divided Fe and FeO, contained in rectangular perforated steel packets instead of tubes.

One of the generally accepted equations for the chemical reaction is as follows:

$$\begin{array}{c} \text{Charged} \qquad\qquad\qquad\qquad\qquad \text{Discharged} \\ \overset{+}{\mathrm{NiO_2}} + (2\mathrm{KOH} + \mathrm{H_2O}) + \overset{-}{\mathrm{Fe}} \rightleftarrows \mathrm{NiO} + (\mathrm{H_2O} + 2\mathrm{KOH}) + \mathrm{FeO}. \end{array}$$

The specific gravity of the electrolyte is about 1.2 at 60°F. This remains sensibly constant, since the solution undergoes no permanent change during either charge or discharge. Hence, **with this type of cell a hydrometer gives no information as to the completeness of charging.**

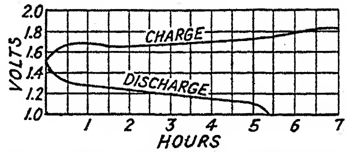


FIG. 384. Charge and Discharge Curves of Edison Cell

The average emf of the Edison cell is 1.2 volts when discharging at the normal rate. The internal resistance depends upon the size of the cell and the state of charge. For a typical 150-ampere-hour cell it is from 0.003 to 0.005 ohm, depending upon the condition of charge. The energy efficiency is 55 to 60%.

Typical curves for the charge and discharge at normal rate of types A, B, and C cells are shown in Fig. 384.

432. The ampere-hour. It is customary to rate the ability of a secondary cell to store up energy in ampere-hours. The *ampere-hour* is defined as the quantity of electricity that flows past a given section of a conductor when a current of one ampere is maintained for one hour.

That is,

$$\begin{aligned} 1 \text{ ampere-hour} &= 1 \text{ ampere} \times 1 \text{ hour} \\ &= 1 \text{ ampere} \times 60 \text{ minutes} \\ &\qquad\qquad = 60 \quad \text{ampere-minutes} \\ &= 1 \text{ ampere} \times 3600 \text{ seconds} \\ &\qquad\qquad = 3600 \ \text{ampere-seconds} \\ &= 3600 \text{ coulombs.} \end{aligned} \tag{349}$$

433. Effect of size of a cell. The emf of a cell depends only upon the chemical nature of its electrodes and electrolyte, and is entirely independent of the size of the cell. When new, the tiny dry cell (Fig. 376) and the large No. 6 cell used for bell ringing have the same emf. The large cell, however, has many times as much energy stored in it as has the smaller cell, and therefore it will last correspondingly longer for a given purpose than will the smaller cell.

Furthermore, the internal resistance, in a general way, varies directly as the distance between the plates and inversely with their effective areas. Hence the values of internal resistances stated in the foregoing paragraphs are only approximately correct for the size of each cell that is most commonly used.

434. Analogy between cells and pumps. In many respects a battery cell or other electric generator causing electricity to flow

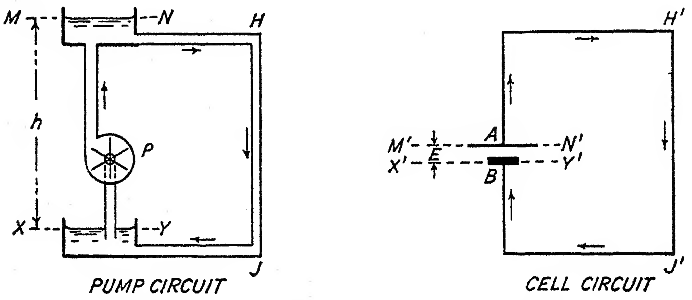


FIG. 385. Analogy Between Cells and Pumps

around a circuit is quite analogous to a pump circulating water around a closed pipe system (Fig. 385).

The pump P does work in elevating water from the level XY to the level MN, from which it flows back downhill to the level XY again.

The cell transfers positive electricity * from the plate B to the plate A, thus raising its potential; and the electricity flows back "downhill" again to B.

On each 1 lb of water that the pump elevates h ft to the upper level, it does h foot-pounds of work. This is the work done on the unit quantity (1 lb) of water by the pump and is therefore the "gravitational potential difference" between the levels MN and XY.

Just so, the cell does an amount of work (V joules) on each unit positive charge (1 coulomb) in transferring it "uphill" from B to A; and V is the electrical potential difference between A and B, as defined in Sec. 353.

As a pound of water flows from the level MN through the pipe line HJ to the level XY, the potential energy h foot-pounds is

* Actually, of course, electrons are caused to flow the other way around.

expended in overcoming friction in the pipe and is dissipated as heat.

Similarly, as a coulomb of electricity flows from A through the circuit $H'J'$ to B, its potential energy (V joules) is expended in overcoming the resistance of the line and is dissipated as heat.

The pump does some work in overcoming the friction of the water in the pump itself for each pound elevated; and the cell does work in overcoming its own internal resistance for each coulomb transferred to A. This internal work of the pump per pound of water added to the gravitational potential difference between MN and XY might be called the "mechano-motive force" of the pump, and represents the total work done by the pump per unit quantity of water elevated. Similarly, the internal work per coulomb done in the cell added to the potential difference between B and A represents the total work done by the cell per unit charge, and this we have called the "electromotive force" of the cell (Sec. 401).

435. Mechanical analogies of potential. From the foregoing paragraph, the "head" h of water is seen to be analogous to potential, for it is the work done per unit quantity (1 lb) of water.

From this, let us compute the work done on another unit quantity (1 ft^3) of water.

By definition, the number of lb in 1 ft^3 is the density D. Hence, when 1 ft^3 of water is elevated h ft, the work done is:

$$W = h \frac{\text{foot-pounds}}{\cancel{\text{pound}}} \times D \frac{\cancel{\text{pounds}}}{\text{ft}^3} = hD \frac{\text{foot-pounds}}{\text{ft}^3}.$$

But by Eq. (109), $hD = p$, the pressure in proper units. Therefore,

$$W = p \frac{\text{foot-pounds}}{\text{ft}^3} = p \frac{\text{pounds}}{\text{ft}^2}.$$

That is, the pressure in proper units represents the work done by a pump per cubic foot of water.

Hence "head of water" is analogous to electric potential when the unit quantity is the pound or the gram; and pressure is analogous to potential when the unit quantity is the cubic foot or the cubic centimeter.

436. Single electrode potentials. Half-cells. When a metal is dipped into a solution containing ions of that metal, e.g., copper in a water solution of $CuSO_4$, there is a potential difference (emf) between the metal and the solution. Nernst, in 1889, undertook to account for this potential difference as follows.

At the surface of the metal there are two tendencies:

1. Metallic ions are tending to leave the solution and deposit on the metal. This tendency is greater, the greater the number of those ions in the solution; i.e., it is proportional to the **ion concentration,** and in dilute solutions this has been shown by Van't Hoff to be approximately proportional to the **osmotic pressure.**

2. There is a tendency also for ions to leave the metal and pass into the solution. Nernst called this **solution pressure** since it can be made to balance the osmotic pressure of the ions by adjusting the concentration.

Assuming that the potential difference E_1 between the electrode and the solution represents the work done on the electric charges when ions pass from the region where the solution pressure is P to the region where the osmotic pressure is p, Nernst derived the formula: *

$$E_1 = -\frac{RT}{nF} \log_e \frac{P}{p} \tag{350}$$

where

R is the general gas constant (8.32 joules per °C per mole);
T is the absolute temperature (273 + °C);
n is the valence of the ion; and
F is the faraday (96,500 coulombs).

Since p is approximately proportional to the concentration C of the ions of the electrode in the solution, Eq. (350) reduces to:

$$\begin{aligned} E_1 &= -\frac{RT}{nF} \log_e \frac{K}{C} \\ &= E_0 - \frac{0.0591}{n} \log_{10} \frac{1}{C} \quad \text{at } 25°\text{C} \end{aligned} \tag{351}$$

where E_0 is the **normal potential** of the electrode; i.e., its potential with respect to the solution when the latter contains 1 gram

* Britton, *Hydrogen Ions* (New York, D. Van Nostrand Co., 1929), p. 11.

equivalent of ions of the electrode per liter of solution, and K is a constant characteristic of the metal electrode.

An electrode in a solution constitutes a half-cell. In general, we cannot connect to the liquid in order to measure the potential difference (emf) of the half-cell without introducing into the solution at least a wire, and this would be another electrode which also would have a potential with respect to the solution. The emf measured would consequently be the algebraic sum of the potential differences between the solution and the two electrodes, respectively.

437. The normal hydrogen electrode. To avoid the difficulty mentioned above, a hydrogen electrode in a normal solution of hydrogen ions has been adopted by international agreement as a standard half-cell whose emf is arbitrarily taken to be zero at all temperatures. This half-cell is often called the *normal hydrogen electrode.*

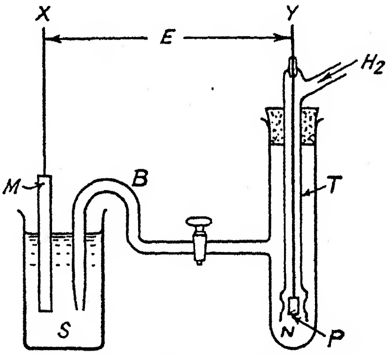


FIG. 386. Normal Hydrogen Electrode

The much-used Hildebrand type is shown at T in Fig. 386. It consists of a thin platinum plate P coated with platinum black and immersed in a normal solution of hydrogen ions N. Hydrogen from a tank or generator passes down the annular glass tube T and some is adsorbed on the platinum black, the excess escaping through holes in the tube. This layer of adsorbed hydrogen is conducting and functions as a plate of hydrogen in a normal solution of hydrogen ions.

To determine the potential difference between an electrode M and a solution S, the latter is connected to the hydrogen half-cell by a capillary "liquid bridge" B, and the emf of the complete cell is measured by means of a potentiometer connected at X and Y. Since the emf of the normal hydrogen electrode is taken to be zero, the potentiometer reading is the required potential difference (emf) between the electrode M and the solution S.

This in principle is the method of finding the potential differ-

ence between each metal and a normal solution of its own ions, as shown in the electromotive series of Sec. 428. Practically, however, it is difficult to maintain a hydrogen electrode accurately normal; hence the more stable calomel electrode is generally used in actual work.

438. The standard calomel electrode, or half-cell. This electrode, or half-cell (Fig. 387), consists of a mass of mercury *A* upon which is a layer of calomel (Hg_2Cl_2), covered by a solution of potassium chloride (KCl) of definite normality and thoroughly saturated with Hg_2Cl_2.

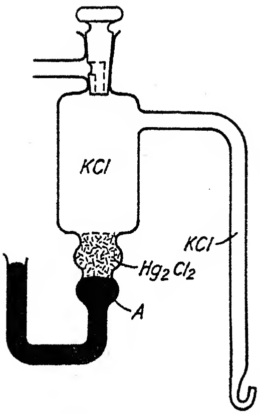


FIG. 387. Standard Calomel Electrode

The absolute value of the potential difference between the Hg and the KCl in this half-cell is determined by the method of the **dropping electrode** or the method of the **capillary electrometer.** If the KCl solution is normal, the potential difference (emf) of the half-cell is found to be 0.560 at 25°C, the mercury being + to the KCl, and **is very stable.**

This type of standard electrode, or half-cell, is therefore generally used in preference to the normal hydrogen electrode for actual determinations.

439. Hydrogen-ion concentration and pH. The acidity or the alkalinity of a solution is expressed numerically by the concentration of hydrogen ions in the solution. Pure water ("conductivity water"), which is neutral, is found by conductivity tests to have a concentration of 10^{-7} gram-ion of H and 10^{-7} gram-radical of OH per liter at 22°C.

Hence, **a neutral solution has a hydrogen-ion concentration of 10^{-7} gram-ion/liter.**

Taking the atomic mass of hydrogen to be 1 for simplicity, a 0.01 normal solution of HCl has 0.01 gram of H-ions per liter and therefore a hydrogen-ion concentration of $0.01 = 10^{-2}$.

In 1909, Sörensen suggested that acidity might be more conveniently represented by merely stating the negative of the exponent of the H-ion concentration, which he designated as its pH value. **The pH value of a solution is the negative of the**

exponent of its hydrogen-ion concentration when that is expressed as a power of 10.

Hence a neutral solution has pH = 7;
a 0.01 normal HCl solution has pH = 2 (more accurately, 2.02); and
a 0.1 normal HCl solution has pH = 1 (more accurately, 1.07).

From this it is seen that increasing acidity is indicated by decreasing pH; and, conversely, increasing alkalinity is indicated by increasing pH.

The **standard method** of determining hydrogen-ion concentration and pH, by which all other methods are checked, is the electrometric method. This is substantially as follows.

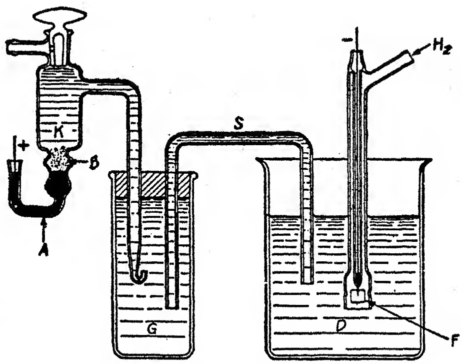


FIG. 388. Assembly for Hydrogen-Ion Determination. (Courtesy Leeds and Northrup Co.)

Figure 388 shows the general arrangement of the apparatus. AKB at the left is the standard calomel half-cell. At the right, forming the other half-cell, is a hydrogen electrode F dipping into the solution D whose hydrogen-ion concentration is to be determined. These two half-cells are connected by the "salt bridge" consisting of the vessel G and the siphon S, which are filled with a saturated solution of KCl to minimize potential differences that might otherwise exist at the surface of contact between the solutions of the two half-cells.

The electromotive force E developed by the complete cell, or "gas chain," is best measured by a potentiometer connected to the terminals + and −.

Calling the emf of the calomel electrode E_{Hg}, and that developed between the solution D and the hydrogen electrode, E_1, we have:

$$E = E_{Hg} - E_1 \tag{352}$$

and substituting for E_1 its value from Eq. (351), n being 1 for hydrogen,

$$E = E_{Hg} - \left(E_0 - \frac{0.0591}{1}\log_{10}\frac{1}{C}\right) \quad \text{at } 25°\text{C.}$$

For a *normal* KCl calomel electrode, $E_{Hg} = 0.560$ at 25°C, and for a normal hydrogen electrode, $E_0 = 0.277$ at 25°C.

Therefore,

$$E = 0.560 - 0.277 + 0.0591\left(\log_{10} \frac{1}{C}\right)$$

$$= 0.283 + 0.0591(-\log_{10} C)$$

$$= 0.283 + 0.0591 \text{ pH} \quad \text{at } 25°\text{C}. \tag{353}$$

This equation is a straight line, and is plotted in Fig. 389. Curves are obtained in a similar manner when 0.1 normal and

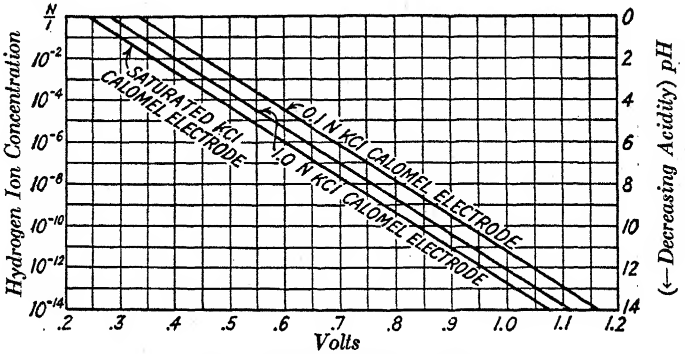


FIG. 389. pH—Emf Curves

saturated KCl calomel electrodes, respectively, are used instead of the normal electrode.

440. International Units. In the preceding pages it will have been observed that, in order to make as many of the constants of proportionality unity as possible, we have defined two sets of

Units of	Name	Definition in emu	Equivalent in esu
Quantity	Coulomb	10^{-1}	3×10^9
Current	Ampere	10^{-1}	3×10^9
Potential	Volt	10^8	$\frac{1}{300}$
Resistance	Ohm	10^9	$\frac{1}{9 \times 10^{11}}$
Inductance	Henry	10^9	$\frac{1}{9 \times 10^{11}}$
Capacitance	Farad	10^{-9}	9×10^{11}

absolute units: the **electrostatic** and the **electromagnetic.** From the latter, at the proper points, we have defined **practical units** as more convenient for commercial purposes. These are tabulated in table on page 520.

Since great skill and care are required to set up the necessary apparatus and to determine the practical units from their definitions, close approximations to their values have been adopted by international agreement as **legal units**, which may be reproduced with comparative ease. These are called **international** units and are defined as follows:

The **international ampere** is that current which in 1 sec will deposit 0.0011180 gm of silver from a solution of silver nitrate, solution and coulombmeter being prepared according to prescribed specifications.

The **international ohm** is the resistance of a uniform column of mercury 106.3 cm long and having a mass of 14.4521 gm at 0°C. (Its cross section is then 1 square millimeter.)

The **international volt** is the potential difference between the ends of a conductor whose resistance is one international ohm when there is a current of one international ampere in the conductor. It is practically 1/1.0183 of the emf of a Weston normal (saturated) cell at 20°C.

From these basic units the others are easily determined.

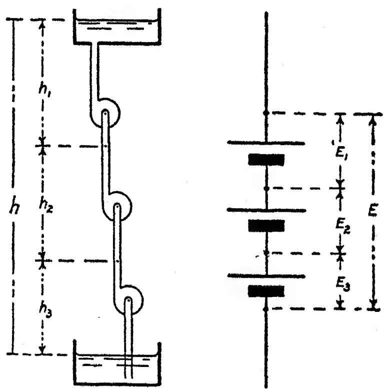


FIG. 390. Cells in Series

441. Cells in series and in parallel. Several cells connected together are called a **battery.** There are two distinct methods of connection: series and parallel. All other methods are combinations of these two.

1. *Series connection.* A group of cells is connected in series when all the electricity in that branch of the circuit passes through all the cells. If the positive terminal of one cell is connected to the negative of the next throughout the group, they are said to be in **conjunction;** otherwise those connected in the opposite way are said to be in **opposition** to the others.

Cells connected in series are analogous to a group of pumps in tandem (Fig. 390). Obviously the total height to which water is pumped, or the total head h due to all the pumps, is:

$$h = h_1 + h_2 + h_3 + \cdots.$$

Similarly, the total emf E due to all the cells of a group in series is:

$$E = E_1 + E_2 + E_3 + \cdots \tag{354}$$

where

E_1, E_2, E_3 are the emfs of the individual cells.

By the law of continuity of current (Sec. 406), all the electricity passes through each cell, and hence,

$$I = I_1 = I_2 = I_3 = \cdots. \tag{355}$$

For the same reason, all the electrons encounter the resistance of each cell, so that the internal resistance r of the group equals the sum of the internal resistances r_1, r_2, r_3, etc. of the individual cells. That is,

$$r = r_1 + r_2 + r_3 + \cdots. \tag{356}$$

2. *Parallel connection.* A group of cells is said to be connected in parallel, or multiple, when only a part of the electrons that

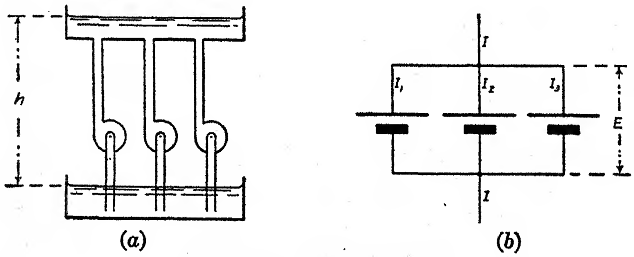


Fig. 391. Cells in Parallel

make up the current pass through each cell. The connection is usually made by joining all the positive terminals together, and all the negative terminals together.

We consider here only the case in which all the cells are exactly alike.*

* Problems involving cells in parallel that are not all alike are best solved by means of Kirchhoff's laws (see Sec. 442).

From the pump analogy of Fig. 391a, the total head h is the same whether all the pumps are working or only one. The difference is that more water is delivered per second when all are working.

Likewise, when all the cells are **exactly alike**, the emf E produced by all is exactly the same as that of a single cell (E_1, E_2, E_3, etc.). That is,

$$E = E_1 = E_2 = E_3 = \cdots. \tag{357}$$

The electrons divide among the several cells, only a part of their total number passing through each cell. The total current I in the main circuit is therefore, by Sec. 406,

$$I = I_1 + I_2 + I_3 + \cdots \tag{358}$$

where I_1, I_2, I_3, etc., are the currents through the individual cells. And since the cells are all exactly alike,

$$I_1 = I_2 = I_3 = \cdots$$

Therefore,

$$I = nI_1 \tag{359}$$

where n is the number of cells in parallel.

The combined internal resistance r of the group is obtained like that for any group of conductors in parallel, by Sec. 414.

$$\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} + \cdots.$$

But the cells being all alike, $r_1 = r_2 = r_3 = \cdots$.

Hence, when n is the number of cells in parallel

$$\frac{1}{r} = \frac{n}{r_1}.$$

Therefore,

$$r = \frac{r_1}{n}. \tag{360}$$

Solution of Simple Circuits: Solved Problems

1. Given the circuit shown in Fig. 392; required the current in each branch and the fall of potential in each section.

Solution: Since the cells are in series, the *total emf* is:

$$E = 1.5 + 1.5 = 3 \text{ volts}.$$

Resistances:

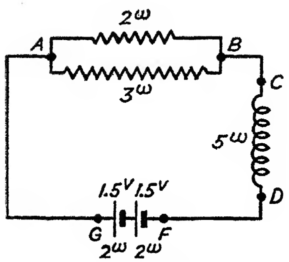


FIG. 392

From A to B, $\frac{1}{R_{AB}} = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$

$R_{AB} = \frac{6}{5}$ ohms.*

From C to D, $R_{CD} = 5$ ohms.

From F to G, $R_{FG} = 2 + 2 = 4$ ohms.

Total resistance † of circuit is:

$$R = \tfrac{6}{5} + 5 + 4 = \tfrac{51}{5} \text{ ohms.}$$

Then by Ohm's law, total current in main circuit is

$$I = \frac{E}{R} = \frac{3}{\frac{51}{5}} = \frac{5}{17} \text{ amp.}$$

Fall of potential in sections of circuit:

ri-drop from A to B, $V_{AB} = \frac{6}{5} \times \frac{5}{17} = \frac{6}{17}$ volt
ri-drop from C to D, $V_{CD} = 5 \times \frac{5}{17} = \frac{25}{17}$ volts
ri-drop from F to G, $V_{FG} = 4 \times \frac{5}{17} = \frac{20}{17}$ volts.

Total ri-drop around circuit $\frac{51}{17} = 3$ volts (check).†

To find currents in 2-ohm and 3-ohm branches:

From above, ri-drop from A to $B \equiv V_{AB} = \frac{6}{17}$ volt

$$\text{Current in 2-ohm branch, } I_2 = \frac{V_{AB}}{2} = \frac{\frac{6}{17}}{2} = \frac{3}{17} \text{ amp}$$

$$\text{Current in 3-ohm branch, } I_3 = \frac{V_{AB}}{3} = \frac{\frac{6}{17}}{3} = \frac{2}{17} \text{ amp.}$$

Adding, total current $I = \frac{5}{17}$ amp (check).

2. Given the circuit shown in Fig. 393; required the current in each branch and the fall of potential in each section.

Solution: Since in each branch of the battery group the cells are in series, the emf of each branch is:

$$E_1 = 2 + 2 = 4 \text{ volts}$$

and since the branches are in parallel, we have:

Total emf of circuit is $E = E_1$
$= 4$ volts.

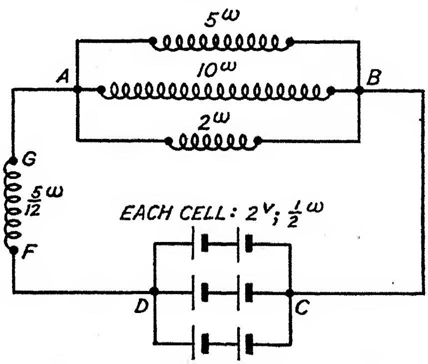


FIG. 393

* If common fractions are retained, results will check exactly, whereas if decimal fractions are used there will be slight "non-checks" due to approximating.

† In circuit problems it is customary to consider the resistance of connecting wires negligible unless it is explicitly given.

Resistances:

From A to B,

$$\frac{1}{R_{AB}} = \frac{1}{5} + \frac{1}{10} + \frac{1}{2} = \frac{8}{10}$$

$$R_{AB} = \tfrac{10}{8} \text{ ohms.}$$

From C to D,

$$\text{resistance in each branch} = \tfrac{1}{2} + \tfrac{1}{2} = 1 \text{ ohm.}$$

Since the three branches are in parallel,

$$\frac{1}{R_{CD}} = \frac{1}{1} + \frac{1}{1} + \frac{1}{1} = \frac{3}{1}$$

$$R_{CD} = \tfrac{1}{3} \text{ ohm.}$$

From F to G,

$$R_{FG} = \tfrac{5}{12} \text{ ohm.}$$

Hence,

Total resistance of circuit, $R = \frac{10}{8} + \frac{1}{3} + \frac{5}{12} = \frac{48}{24} = 2$ ohms.

Then by Ohm's law, *total current* in main circuit is:

$$I = \frac{E}{R} = \frac{4}{2} = 2 \text{ amp.}$$

Fall of potential in sections of circuit:

$$\begin{aligned}
\text{ri-drop from } A \text{ to } B,\ V_{AB} &= \tfrac{10}{8} \times 2 = \tfrac{5}{2} \text{ volts}\\
\text{ri-drop from } C \text{ to } D,\ V_{CD} &= \tfrac{1}{3} \times 2 = \tfrac{2}{3} \text{ volt}\\
\text{ri-drop from } F \text{ to } G,\ V_{FG} &= \tfrac{5}{12} \times 2 = \tfrac{5}{6} \text{ volt}\\
\text{Adding, total ri-drop around circuit} &= \tfrac{24}{6} = 4 \text{ volts (check).}
\end{aligned}$$

Currents in branches of divided sections:

$$\begin{aligned}
\text{In 5-ohm branch, } I_5 &= \frac{V_{AB}}{5} = \frac{\frac{5}{2}}{5} = \frac{1}{2} \text{ amp}\\
\text{In 10-ohm branch, } I_{10} &= \frac{V_{AB}}{10} = \frac{\frac{5}{2}}{10} = \frac{1}{4} \text{ amp}\\
\text{In 2-ohm branch, } I_2 &= \frac{V_{AB}}{2} = \frac{\frac{5}{2}}{2} = \frac{5}{4} \text{ amp}\\
\text{Adding, total current } I &= \tfrac{8}{4} = 2 \text{ amp (check)}\\
\text{In 1 branch of } CD,\ I_1 &= \frac{V_{CD}}{1} = \frac{\frac{2}{3}}{1} = \frac{2}{3} \text{ amp.}
\end{aligned}$$

Since all three branches of battery circuit are alike, the current is the same in each, i.e., 2/3 ampere; hence

$$\text{Total current } I \text{ is } 3 \times \tfrac{2}{3} = 2 \text{ amp (check).}$$

442. Kirchhoff's laws. In the preceding problems it was seen that if an electric circuit can be resolved into an equivalent simple series circuit, it may be solved by means of Ohm's law. But this

is not possible in a complex network of conductors and generators such as a street railway system.

From Ohm's law and the law of continuity of current (Sec. 406), G. R. Kirchhoff (1824–87) formulated two laws **by means of which any electrical network may be solved.** They may be derived as follows.

After the switches have been closed and the current has become constant in all branches of a given network (Fig. 394), as much electricity must flow away from any junction A in one second as flows toward it, for otherwise the law of continuity of current would be violated. This is equivalent to Kirchhoff's first law:

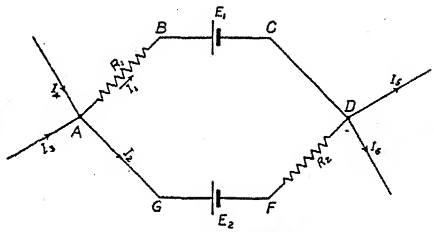


FIG. 394. Kirchhoff's Law

The algebraic sum of the currents at any junction must be zero. The usual convention of signs is to take currents toward a junction +, and those away, −.

Expressing this in symbols, at A

$$I_3 - I_1 - I_2 + I_4 = 0$$

or

$$\Sigma I = 0. \tag{361}$$

Consider next any closed circuit $ABCDFGA$ in any network. From all sources of emf, shown and not shown, there will be a definite potential V_A at A and V_D at D; and the potential difference between A and D will be the definite quantity $(V_A - V_D)$, regardless of the path by which one passes from A to D.

Along the upper path, the potential difference $(V_A - V_D)$ is opposed by the emf E_1. Applying Ohm's law to this branch of the network,

$$\begin{aligned}(V_A - V_D) - E_1 &= R_1 I_1 \\ (V_A - V_D) &= E_1 + R_1 I_1. \end{aligned} \tag{a}$$

Similarly for the lower path from A to C,

$$\begin{aligned}(V_A - V_D) - E_2 &= R_2 I_2 \\ (V_A - V_D) &= E_2 + R_2 I_2. \end{aligned} \tag{b}$$

From Eqs. (b) and (a),

$$E_2 + R_2I_2 = E_1 + R_1I_1$$

whence

$$E_2 - E_1 = R_1I_1 - R_2I_2$$

or

$$\Sigma E = \Sigma RI. \tag{362}$$

Kirchhoff's second law is represented by Eq. (362). In words it is: **The algebraic sum of the emfs around any closed circuit is equal to the algebraic sum of the ri-drops.** For the law in this form, the following rules for signs are found by inspection of the figure.

Trace around any closed circuit always in the same sense. When a generator is traced through from its negative to its positive terminal, its emf is to be taken +, otherwise −. When a resistance is traced through in the same sense as the current, the ri-drop is to be taken +, otherwise −.

Solved Problems

1. Let it be required to find the current in each branch of the network of Fig. 395a, and the potential difference between A and D.

Applying Kirchhoff's first law at A,

$$I_2 - I_1 + I_3 = 0. \tag{a}$$

Applying Kirchhoff's second law around circuit $ABCDA$,

$$-3 + 8 = 3I_1 + 12I_2$$
$$5 = 3I_1 + 12I_2 \tag{b}$$

and around circuit $ADFHA$,

$$-8 + 6 = -12I_2 + .5I_3$$
$$-2 = -12I_2 + .5I_3. \tag{c}$$

Solving Eqs. (a), (b), and (c) simultaneously, we get:

$$I_1 = \tfrac{77}{87} \text{ amp}$$
$$I_2 = \tfrac{17}{87} \text{ amp}$$
$$I_3 = \tfrac{60}{87} \text{ amp}.$$

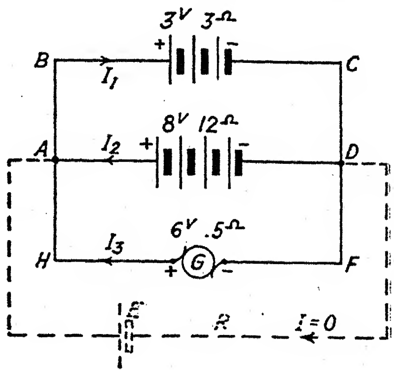


FIG. 395a

In order to find the potential difference between the two points A and D, imagine a fictitious circuit (shown dotted) between A and D, containing a battery whose emf E is *exactly equal and opposite* to the potential difference between A and D. Then, regardless of the resistance of this fictitious circuit, the current in it will be zero.

Applying Kirchhoff's second law around the fictitious circuit and $ABCD$:

$$E - 3 = 3I_1 + R \times 0$$
$$E = 3 + 3(\tfrac{77}{87}) = 5\tfrac{19}{29} \text{ volts.}$$

This value may be checked by tracing through the fictitious circuit and any other branch. Since E is equal and opposed to the potential difference between A and D, A must be + to D.

2. In the network shown in Fig. 395b given $I_1 = 2$ amp., $I_2 = 3$ amp., $R_1 = 5\Omega$, $R_2 = 4\Omega$, $R_3 = 10\Omega$, $R_5 = 2\Omega$. Required I_3, I_4, I_5, R_4.

Applying Kirchhoff's first law,

$$\text{at B } I_1 + I_5 - I_3 = 0$$
$$2 + I_5 - I_3 = 0 \qquad \text{(a)}$$
$$\text{at D } I_2 - I_5 - I_4 = 0$$
$$3 - I_5 - I_4 = 0. \qquad \text{(b)}$$

Applying Kirchhoff's second law, around circuit $ABDA$

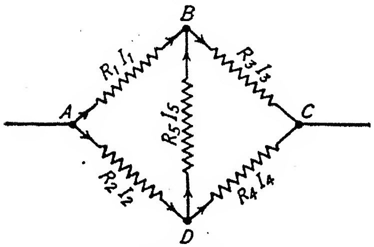


Fig. 395b

$$0 = R_1I_1 - R_5I_5 - R_2I_2$$
$$0 = 5 \times 2 - 2I_5 - 4 \times 3 \qquad \text{(c)}$$
$$I_5 = -1 \text{ amp.}$$

The (−) sign means that the wrong direction was assumed for I_5. The magnitude, however, is correct and the sign **should not be changed** unless the problem is reworked from the beginning.

Around circuit $ABCDA$

$$0 = R_1I_1 + R_3I_3 - R_4I_4 - R_2I_2$$
$$0 = 5 \times 2 + 10I_3 - R_4I_4 - 4 \times 3. \qquad \text{(d)}$$

Substituting the value of I_5 in (a),

$$2 + (-1) - I_3 = 0$$
$$I_3 = 1 \text{ amp.}$$

Substituting the value of I_5 in (b),

$$3 - (-1) - I_4 = 0$$
$$I_4 = 4 \text{ amp.}$$

Substituting the values of I_3 and I_4 in (d),

$$0 = 10 + 10 \times 1 - R_4\,4 - 12$$
$$R_4 = 2 \text{ ohms.}$$

PROBLEMS

1. If the electrochemical equivalent of copper is 0.000329 gm/clmb, how long will it take a current of 0.5 amp to deposit 1.8 gm on the cathode?

2. If 0.001118 gm of silver is deposited by 0.2 amp in 5 sec, what mass of zinc would be deposited from a zinc sulphate solution by the same current in that time?

3. If the electrochemical equivalent of copper is 0.000329 gm/clmb, how long will it take a current of 0.6 amp to deposit 2 gm of copper?

4. If 0.438 gm of nickel is deposited by a current of 0.6 amp in 40 min, compute the electrochemical equivalent of nickel.

5. If 0.006708 gm of silver is liberated by 2 amp in 3 sec, what mass of copper would be deposited in the same circuit?

6. How many grams of water would be decomposed by a current of 8 amp in 1.5 hr?

7. How many cubic centimeters of water would be decomposed by a current of 5 amp in 2 hr?

8. If the density of hydrogen gas is 0.000090 gm/cm³, how long will it take a current of 5 amp to produce 1 liter of hydrogen by electrolysis of water?

9. If the atomic mass of zinc is 65.38 and its valence is 2, compute the electrochemical equivalent of zinc from that of silver.

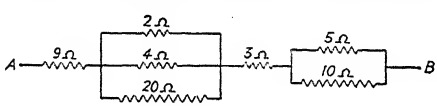
10. A Daniell cell having an emf of 1.08 volts is connected to the terminals of a resistance of 2.4 ohms and produces a current of 0.15 amp. What is the internal resistance of the cell?

11. A dry cell whose emf is 1.5 volts is connected to the terminals of a resistance of 3.2 ohms and produces a current of 0.25 amp. What is the internal resistance of the cell?

12. A lead storage cell has an emf of 2.1 volts, but when producing a current of 15 amp its terminal voltage is only 2.06 volts. What is the internal resistance of the cell?

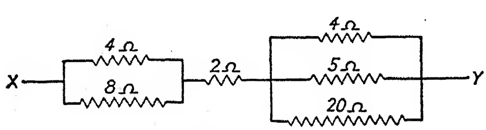
13. How many cells in series, each having an emf of 1.3 volts and an internal resistance of 0.5 ohm, will be needed to produce a current of 1.5 amp in an external resistance of 20 ohms?

14. When two cells that are just alike are connected in series they produce a current of 0.25 amp in an external resistance of 8 ohms. When connected in parallel they produce 0.16 amp in the same resistance. Find the emf and the internal resistance of each cell.



PROB. 16

15. How many cells in series, each having an emf of 1.2 volts and an internal resistance of 0.4 ohm, will be needed to produce a current of 2 amp in an external resistance of 30 ohms?



PROB. 17

16. Find the total resistance between A and B.

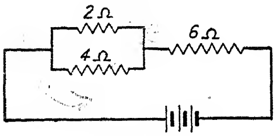
17. Find the total resistance between X and Y.

18. When resistances of 2, 3, and 5 ohms, respectively, are joined in parallel and a total current of 4 amp is produced in the system, what is the current in each resistance?

19. Three coils of 50, 75, and 15 ohms, respectively, are joined in parallel

and a total current of 3 amp is produced in the system. What is the current in each coil?

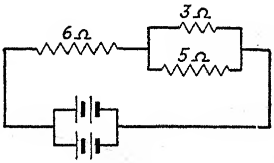
20. Resistances of 25, 40, and 60 ohms are connected in parallel and a current of 5 amp is produced in the system. What is the current in each resistance?



PROB. 21

21. If each cell in the diagram has an emf of 1.5 volts and an internal resistance of 1.0 ohm, find the current in the 4-ohm branch.

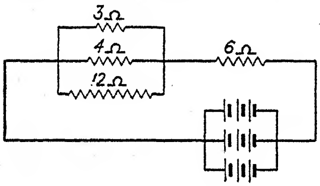
22. In a circuit similar to the above, the coils in parallel have resistances of 10 and 20 ohms, respectively, and the series coil, 5 ohms. If the emf of each cell is 1.25 volts and its internal resistance is 2 ohms, find the current in each coil.



PROB. 23

23. Find the current in each branch of the circuit shown. Each cell has an emf of 1.5 volts and an internal resistance of 2.4 ohms.

24. In a circuit similar to the above the coils in parallel have resistances of 2 and 4 ohms, respectively, and the series coil, 3 ohms. If each cell has an emf of 1.6 volts and an internal resistance of 2.2 ohms, find the current in each coil.



PROB. 25

25. Each cell in the circuit shown has an emf of 1.4 volts and an internal resistance of 0.6 ohm. Find the current in the 4-ohm coil.

26. In a circuit similar to the above, the coils in parallel have resistances of 2, 4, and 6 ohms, respectively, and the series coil, 5 ohms. If each cell has an emf of 1.5 volts and an internal resistance of 0.5 ohm, find the current in the 4-ohm coil.

CHAPTER XXX

THERMOELECTRICITY

443. Contact potential difference. Alessandra Volta, who devised the first battery cell, also discovered that if two different metals are placed in contact, a potential difference will in general develop between them, which will depend upon the temperature and the nature of the metals. For example, when a piece of iron is in contact with a piece of copper, the iron exhibits a + charge and the copper a − charge, the potential difference being 0.15 volt at 20°C.

Contact differences of potential are explained by the fact that the free electrons in a metal behave collectively much like a gas. In general, the density of this "electron gas" is not the same in different metals. Hence, if two pieces of different metals are placed in contact, electrons will diffuse from the one of higher electron density to the one of lower, until the opposing field that builds up is sufficient to stop the flow. An emf is thus developed which internally is measured from negative to positive, since that is the direction of the current it would produce if the circuit were completed externally.

444. The thermoelectric effect. While studying such contact potential differences, J. T. Seebeck at Berlin, about 1821, noted that if a circuit is composed of two wires of different metals, and the two junctions are kept at different temperatures, there will generally be a current of electricity in the circuit. This phenomenon is known as the **thermoelectric, or Seebeck, effect.**

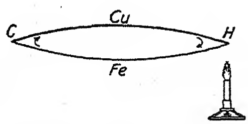


FIG. 396. Thermocouple

Thus, if copper and iron wires are used to form the circuit of Fig. 396, and the junction H is heated to a temperature higher than that of the junction C, the current will be from copper to iron at the hot junction.

Let a high-resistance millivoltmeter be introduced into the

copper-iron circuit at some point (Fig. 397); and let the cold junction be maintained at 0°C while the temperature of the hot junction is raised. It will be found that the emf developed in the circuit will increase, becoming a maximum when the temperature of the hot junction reaches 275°C; then it will decrease, becoming zero again when the temperature of the hot junction is 550°C. For still higher temperatures of H the emf is reversed, producing a current from iron to copper at the hot junction.

Fig. 397

The relation between the temperature of the hot junction and the emf is very nearly parabolic, as shown in Fig. 398. If the cold junction is maintained at any temperature other than 0°C, the result is equivalent to moving the axes to the point on the curve that corresponds to that temperature. Thus, if the cold junction were at 100°C, the curve would be unchanged but the axes would be shifted, as shown by AX and AY.

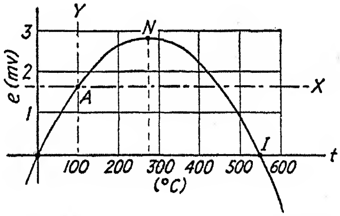


Fig. 398. Curve for Copper-Iron Couple

The temperature at which the tangent is horizontal is called the **neutral temperature**; that at which the emf reverses, the **temperature of inversion.**

For some couples the vertex of the parabola will be upward; for others, downward. Thus, for a platinum-lead couple the neutral point is at −150°C and the curve convex upward; whereas for a cadmium-lead couple the neutral point is at −100°C, and the curve convex downward (Fig. 399).

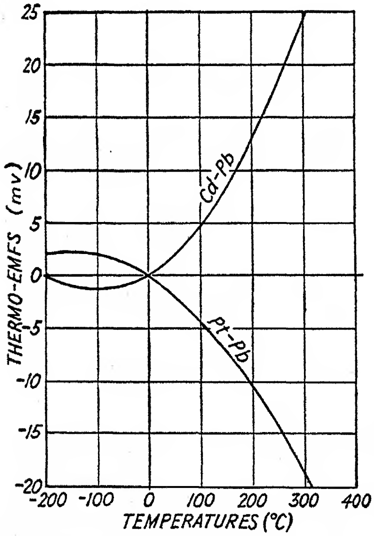


Fig. 399. Temperature-Emf Curves

445. The Peltier effect. About thirteen years after Seebeck's discovery, Peltier, a Parisian watchmaker, discovered the reverse

effect. That is, if the heating device is removed and the junctions C and H are allowed to return to the same temperature, and if a source of emf B is then introduced into the circuit (Fig. 400) so as to produce **a current in the same direction** as before, it will be found that the junction H is now cooled, and the junction C heated by the current from the battery. If the **current is reversed,** H is heated and C is cooled, and the rate of heating and cooling is proportional to the current; whereas the "Joule heat," developed when a current is maintained in a resistance, is proportional to the square of the current (Sec. 419).

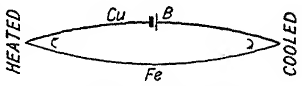


FIG. 400. Peltier Effect

This would indicate that there is an emf (called a "Peltier emf") at each junction of the two metals. At H the emf is with the current produced by the cell, when connected as shown in Fig. 400. The electron stream therefore gains energy as it passes through this junction, at the expense of the heat of the junction, and the junction is consequently cooled. At C just the reverse takes place: the junction emf is opposed to the cell current. Work is therefore done against this emf in order to maintain the current, and a corresponding amount of heat is liberated.

Such a circuit, or thermocouple, is therefore a heat engine (Sec. 327), taking in heat at one junction, converting some of it into electrical energy, and discharging the remainder at the other junction.

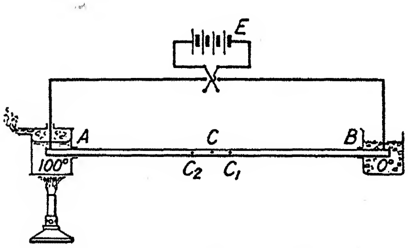


FIG. 401. Thomson Effect

446. The Thomson effect. In studying the changes with temperature of the emf of a thermocouple, Lord Kelvin (William Thomson) concluded that there must be other sources of emf in the circuit besides those at the junctions. Experiment confirmed this. If a piece of copper wire AB (Fig. 401), sufficiently large to make its resistance negligible, has its ends maintained at different temperatures, say, 100°C and 0°C, there will be a point C on the wire at which the temperature is 50°C.

If the terminals of the wire are then connected to a battery E,

so that a current is produced from A to B, the 50° point will move to C_1 toward the cold end B; i.e., the wire as a whole is heated.*

Since the resistance is negligible, there is no "Joule heating"; hence the heating must be due to the fact that work is being done in pumping the electricity against an emf. On reversing the current, the 50° point is shifted to C_2 toward the hot end; i.e., the wire as a whole is cooled. This is what would be expected if the emf in the wire is doing work in maintaining the current. Moreover, the rate of heating and cooling is proportional to the first power of the current, as in the case of the Peltier emf's at the junctions.

The phenomenon described above is known as the *Thomson effect*, and it shows that an emf is developed in a conductor if different parts are not at the same temperature. The occurrence of these "Thomson emf's" may be accounted for as follows.

Since the electrons in a conductor behave somewhat like a gas, we should expect them to be more concentrated in cold metal than in hot metal. Hence the cold end should be negative to the hot end; i.e., there should be an emf from the cold to the hot end, since this is the direction of the current when an external circuit is completed. This is found to be true in copper; but in iron and in some other metals just the reverse is found, which as yet is unexplained. The Thomson effect does not occur in lead; i.e., Thomson emf's in lead are zero. For this reason, lead is the metal with reference to which the thermo-emf's of other metals are usually given.

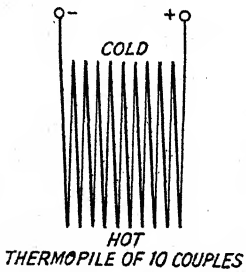


FIG. 402. Thermopile

447. Thermocouples. A thermocouple (Fig. 396) is a circuit of two different metals: it converts heat directly into electrical energy when the two junctions are at different temperatures. Several thermocouples connected in series are called a **thermopile** (Fig. 402).

The electromotive force developed by a thermocouple is the algebraic sum of the Peltier emf's at the junctions and the Thom-

* This is not a satisfactory demonstration experiment. For a better method see Starling, *Electricity and Magnetism* (London, Longmans, Green and Company, 1937), p. 208.

son emf's in the wires. The order of magnitude of thermo-emf's is a few hundredths of a volt; hence the use of thermocouples is largely confined to temperature determinations, for which in many instances they have no equal.

Analysis of the relation of thermo-emf to temperature yields the following laws:*

1. Thermo-emf's are additive. That is, if a thermocouple develops an emf e_{12} when the cold junction is at t_1 and the hot at t_2, and e_{23} when the cold is at t_2 and the hot at t_3; then when the cold junction is at t_1 and the hot junction at t_3, the couple will develop an emf e_{13}, such that

$$e_{13} = e_{12} + e_{23}. \tag{363}$$

2. If e_{ac} is the thermo-emf of a metal a with respect to a metal c; and e_{bc} is the thermo-emf of a metal b with respect to metal c; then the emf e_{ab} of metal a with respect to metal b is:

$$e_{ab} = e_{ac} - e_{bc}. \tag{364}$$

3. The emf of a thermocouple will not be altered by the introduction of additional metals into the circuit, provided that **both ends** of each additional metal are kept at the same temperature.

4. The relation between thermo-emf and temperature is approximately parabolic.

These laws have been fully justified by experiment.

In preparing to measure thermo-emf's, the couple is usually opened at the cold junction, and to the two ends thus obtained are soldered copper wires which connect to the potentiometer or high resistance millivoltmeter. The introduction of this copper wire into the circuit does not affect the emf of the couple, as both its ends are kept at the same temperature (see law 3, above).

The two junctions that have been made from the original cold junction are still called the cold junction, and are usually kept (well insulated from each other) in a Dewar flask of melting ice. This is not necessary, however, since the emf can be corrected for temperature of the cold junction by the first law, above.

* Page and Adams, *Principles of Electricity* (New York, D. Van Nostrand Company, 1931), p. 218.

The relations most commonly used are Holborn and Day's equation:

$$e = a + bt + ct^2 \tag{365}$$

and Holman's equation:

$$e = mt^n \tag{366}$$

where e is the thermo-emf; t is the temperature of the hot junction; and a, b, c, and m, n, are constants depending upon the kinds of metals used and the choice of units.

When Holman's equation is used, the temperature of the cold junction must obviously be at zero. Holborn and Day's equation represents experimental results more accurately than does Holman's.

The constants should be determined for each couple just before and after it is used, because couples of the same nominal composition differ among themselves, and a given couple varies with age and use.*

Ranges of couples. The temperature ranges in which the various couples have been found most satisfactory are:

Base Metal Couples:

Copper-constantan	-100 to 500°C
Iron-constantan	-100 to 800
Chromel-alumel	0 to 1100

Noble Metal Couple:

Platinum-platinum-rhodium	0 to 1600

The compositions of these alloys are:

Constantan (eureka, ideal)	60 Cu,	40 Ni
Chromel	10 Cr,	90 Ni
Alumel	98 Al,	2 Ni
Platinum-rhodium	90 Pt,	10 Rd

With the exception of the resistance thermometer, thermocouples are the most sensitive means of measuring temperatures from $-100°$ to 1600°C.

* See *Dictionary of Applied Physics* (London, The Macmillan Company, 1922), I, 917.

448. Vacuum thermocouples. By enclosing a thermocouple in a highly exhausted vessel, loss of heat by conduction and convection is greatly reduced. If also the thermo-elements have low heat conductivity, as in the case of antimony and bismuth, the heat lost by conduction is still further reduced. By these means the sensitivity of thermocouples is enormously increased. Vacuum thermocouples of this type are used for measuring the radiation from stars. Such a couple (Fig. 403), with a high sensitivity galvanometer, will register the radiation from a candle at a distance of 53 miles, provided a telescope of sufficient light gathering power is used to focus the image of the candle upon the thermo-junction.

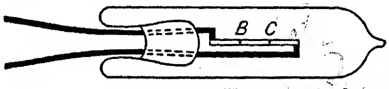


FIG. 403. Vacuum Thermocouple

449. Thermoelectric power. The rate of change of thermoelectromotive force with temperature of the hot junction, i.e., the change of emf per degree change of temperature of the hot junction, is called the *thermoelectric power* of the couple.

Differentiating Eq. (365) with respect to t, we have:

$$P \equiv \frac{de}{dt} = b + 2ct \tag{367}$$

which represents a straight line. For this reason it is somewhat more convenient to plot thermoelectric power curves than thermo-emf curves; hence the former are largely used for comparison purposes. Values of the coefficients b and $2c$ for various substances with respect to lead are to be found in tables of physical constants.

CHAPTER XXXI

Starting Lee

ELECTROMAGNETISM

450. Ampère's law. Magnetic field at center of a short coil. In Sec. 394 we used a simplified form of Ampère's law.* Its more general form is:

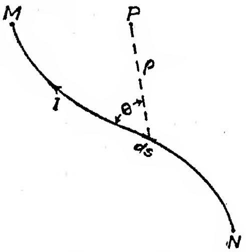


FIG. 404. Ampère's Law

$$dH = \frac{I \sin \theta \, ds}{\rho^2} \tag{368}$$

where dH is the component of the field intensity **at point P** (Fig. 404) due to the current I abamperes in the element of length ds of the conductor, and θ is the angle between the direction of ds and the line joining it to P. The direction of dH is perpendicular to the plane of ds and ρ and its sense into the paper, in this case.

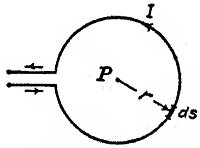


FIG. 405. Current in Circular Coil

Applying this to a circular coil of one turn of radius r cm, in which the current is I abamperes, as in Fig. 405, we have

$$\theta = 90°; \ \sin \theta = 1; \ \rho = r.$$

Hence, by Eq. (368), the field intensity at P is

$$H = \int_0^{2\pi r} \frac{I \sin \theta \, ds}{\rho^2} = \frac{I}{r^2} \int_0^{2\pi r} ds$$

$$= \frac{2\pi I}{r} \text{ oersteds.} \tag{369}$$

If there are N turns and the current is I' amperes, the value of the field intensity becomes:

$$H = \frac{2\pi N I'}{10 r} \text{ oersteds.} \tag{370}$$

* Known also as the law of Laplace, and as the law of Biot and Savart, all of whom did pioneer work in establishing it.

We are now prepared to derive the simplified form of Ampère's law used in Sec. 394.

From the definition of field intensity, a $+1$ pole at P in Fig. 405 is acted upon by the force $\dfrac{2\pi I}{r}$ dynes, according to Eq. (369). By Newton's third law, the $+1$ pole reacts on the current with an equal and opposite force F, so that in magnitude,

$$F = H = \frac{2\pi I}{r} = \frac{1}{r^2}(2\pi r I).$$

But by Sec. 385, $\dfrac{1}{r^2}$ lines/cm² is the radial flux density B produced at the wire by the $+1$ pole; and $2\pi r$ cm is the length l of the current perpendicular to the flux.

Therefore,

$$F = BlI$$

which is the form of Sec. 394.

451. Magnetic fields produced by currents in conductors of various shapes.

1. *Magnetic field about a long straight wire.* Let P be a point at a distance r cm from a very long straight wire in which the current is I abamperes, as shown in Fig. 406. Applying Ampère's law, and noting that

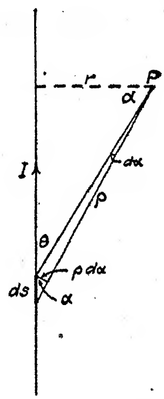


FIG. 406. Current in Straight Wire

$$\sin\theta = \cos\alpha;\ \rho = r\sec\alpha;\ ds = (\rho d\alpha)\sec\alpha = r\sec^2\alpha\, d\alpha,$$

$$dH = \frac{I\sin\theta\, ds}{\rho^2} = \frac{I\cos\alpha\, r\sec^2\alpha\, d\alpha}{r^2\sec^2\alpha} = \frac{I}{r}\cos\alpha\, d\alpha$$

$$H = \frac{I}{r}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos\alpha\, d\alpha = \frac{I}{r}\Big[\sin\alpha\Big]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{2I}{r}\ \text{oersteds.} \tag{371}$$

2. *Work to move a magnet pole around a current.* Consider a $+1$ pole at any distance r cm from the wire of Fig. 406. The force acting upon it is $2I/r$ dynes. To push the pole completely around the circumference will require an amount of work W:

$$W = \frac{2I}{r} \times 2\pi r = 4\pi I\,\frac{\text{ergs}}{\text{unit pole}}. \tag{372}$$

It will be observed that the radius cancels out; hence the work required is independent of the path of the pole around the current.

3. *Magnetic field on center line of a toroidal coil.* Let I abamperes be the current in a uniform toroidal coil of N turns (Fig. 407).

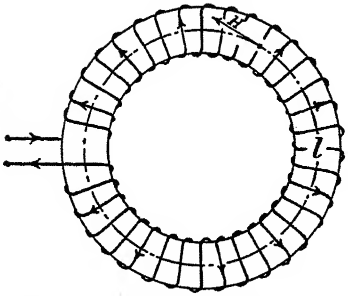


Fig. 407. Field in Toroidal Coil

Assuming perfect symmetry, the field intensity H will be the same at every point of the center line, whose circumference is l.

If a $+1$ pole is pushed around this circumference in opposition to H, the work done will be:

$$W = Hl.$$

During the round trip, the pole will pass around the current in each of the N turns of the coil. Hence, by Eq. (372), the work is also expressed by $N(4\pi I)$. Therefore,

$$Hl = 4\pi NI$$

$$H = \frac{4\pi NI}{l} \text{ oersteds.} \tag{373}$$

When I' is in amperes,

$$H = \frac{4\pi NI'}{10l} \text{ oersteds.} \tag{374}$$

4. *Magnetic field at the center of a long helix, or solenoid.* If the length l of the toroidal coil of the preceding paragraph is great compared to the diameter of one turn, the coil may be straightened out into a long helix, as in Fig. 408, and the field intensity at its center will be unaltered. That is, **at the center** of a long helix,

$$H = \frac{4\pi NI'}{10l} \text{ oersteds} \tag{375}$$

but at the ends,

$$H = \frac{2\pi NI'}{10l} \text{ oersteds} \tag{376}$$

where I' is in amperes in both formulas.*

* For derivation of these and other forms see Culver, *Electricity and Magnetism* (New York, The Macmillan Company, 1939), p. 185.

452. Magnetomotive force. Magnetomotive force $\mathfrak{M}$ is defined as **the work required to carry a +1 pole completely around a magnetic circuit.** Like electromotive force, it is a misnomer, for it is not force but *work per unit pole.*

Let a +1 pole be moved against the field intensity H completely around a closed line of force that interlinks all the turns of any coil in which the current is I abamperes as in Fig. 408.

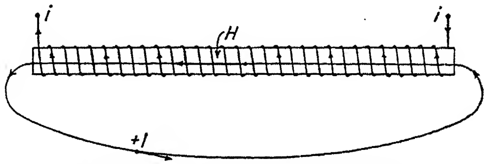


FIG. 408. Current in Long Helix

On the round trip, the pole will pass around the current in each of the N turns of the coil. Hence, by Sec. 451, the work done, i.e., the magnetomotive force of the current in the coil, is

$$\mathfrak{M} = N(4\pi I) \text{ ergs/unit pole, or gilberts.} \tag{377}$$

And if the current is I' amperes,

$$\mathfrak{M} = \frac{4\pi NI'}{10} \text{ gilberts.} \tag{378}$$

453. Units of magnetomotive force. As used above, the gilbert * is defined as the magnetomotive force producing the flux when 1 erg of work is required to carry a +1 pole around the circuit against the field intensity.

But I' amperes and N turns are the physical factors in Eq. (378). Hence it is usual in practice to express magnetomotive force in ampere-turns.

$$NI' \text{ ampere-turns} \equiv \frac{4\pi NI'}{10} \text{ gilberts}$$

$$1 \text{ ampere-turn} \equiv 0.4\pi \text{ gilberts} \tag{379}$$

454. Permeability. While the term "permeability" was introduced in Sec. 379, it was not defined there because we do not determine permeability from Coulomb's law. The point has now been reached where we may define permeability and flux density in terms of the procedure for finding their values.

* Named after Dr. William Gilbert (Sec. 375).

Consider a toroidal coil C of N turns having *no core but air*, the length of its circular center line being l. Let a search coil S of insulated wire be wound over C and connected to a ballistic galvanometer G, as shown in Fig. 409.

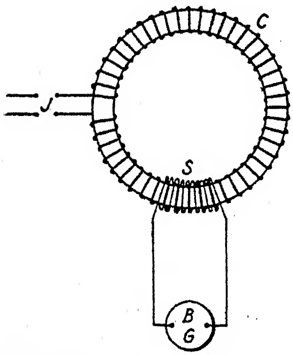


FIG. 409. Determination of Permeability

If a current I is suddenly set up in C by closing switch J, there will be a throw θ_a of the galvanometer.

Let the experiment be repeated, keeping everything the same as before except that the toroidal coil is now wound on an *iron core*. On closing the switch, the current in C will be the same as before, but the galvanometer throw θ_i will be much greater than before.

As the coil S is insulated from coil C, the deflection of G must be due to an emf induced in S when the magnetic flux through it changes, in accordance with the law of Henry and Faraday.

Now, by Eq. (373), the field intensity H, or magnetizing force, in coil C is:

$$H = \frac{4\pi NI}{l} \text{ oersteds.}$$

Since N, I, and l were the same in both cases above, H was the same in both. Hence, to account for the different deflections, we must hypothesize something else (which may involve H) within the coils. That something is what we have called magnetic flux Φ; and flux per unit area we call **flux density *B*.**

The experiment shows that with the same field intensity, a greater flux permeates the iron than the air. Therefore the permeability of iron is said to be greater than that of air.

The permeability μ of a substance is defined as the ratio of the flux density B in the substance to the accompanying field intensity H.

$$\mu \equiv \frac{B}{H} \tag{380}$$

Unit permeability is defined as the permeability of a vacuum,

but the permeability of air (1.0000004) may be taken as unity for practical purposes. Hence, for air,

$$B_a = H.$$

Numerical values of μ and B may be determined as follows. Let θ_a, B_a, and θ_i, B_i, be the throws (or first deflections) and the flux densities for air and iron, respectively. The theory of the ballistic galvanometer shows that

$$\frac{\theta_i}{\theta_a} = \frac{B_i}{B_a} = \frac{B_i}{H} = \mu_i$$

whence

$$B_i = \mu_i H.$$

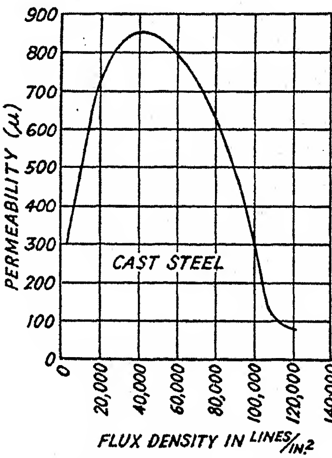


FIG. 409a. Variation of Permeability

454–1. Variation of permeability. The permeability of ferromagnetic substances varies with the flux-density (Fig. 409a) and with the temperature, but not linearly. There is a critical temperature, called the **Curie point,** above which they cease to be ferromagnetic and behave like paramagnetic substances. For iron the Curie point is about 780°C and is identical with one of the points of recalescence.

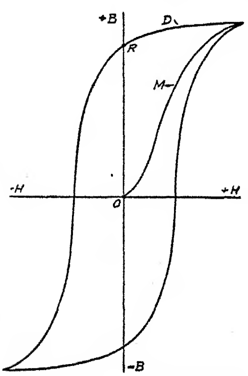


FIG. 409b. Hysteresis Loop

Alexanderson has found that the permeability of iron is practically independent of the frequency of reversal of magnetization.

The permeability of paramagnetic bodies is sensibly constant for different flux densities; i.e., B is directly proportional to H. It varies, however, with temperature.

454–2. Hysteresis. If a piece of ferromagnetic material is subjected to a variable magnetic field, and the values of B plotted against the corresponding values of H, a "magnetization curve" M, Fig. 409b, will be obtained as the intensity increases. When the intensity

H is then decreased, the curve M is not retraced; but a new, demagnetization curve D is obtained. That is, the values of B lag behind those of H; and when H becomes zero again, B has still the value OR called the *residual magnetism*.

This lagging of magnetic induction (B) behind the field intensity (H) is called **magnetic hysteresis.**

If H is carried through a cycle of values, positive and negative, as when the magnetization is produced by an alternating current, the closed curve called a *hysteresis loop* is obtained.

It has been shown by Ewing that the energy required to carry the specimen through the cycle of magnetization is proportional to the area of the hysteresis loop. Such energy is transformed into heat and represents a considerable loss in alternating current machinery.

NOTE. **The Magnetic Circuit** is discussed in the Appendix, page 812.

CHAPTER XXXII

ELECTROMAGNETIC GENERATORS

455. Electromagnets. In Sec. 393, the fact was noted that a coil of wire in which a current of electricity is maintained has all the properties of a magnet. The effect is greatly increased if a core of soft iron is placed in the coil, as was found by Dominique François Arago in 1820. The device is then called an *electromagnet* (Fig. 410).

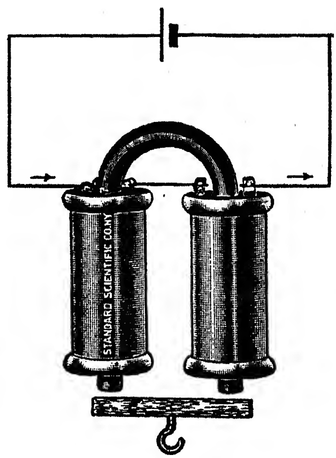


FIG. 410. Electromagnet. (Courtesy Chicago Apparatus Co.)

Since soft iron loses almost all of its magnetism promptly when the current ceases, magnets of this kind are essential elements of the telegraph sounder, the telephone, electric bells, etc. These applications will be discussed in a later chapter.

456. The simple alternator. The law of Henry and Faraday suggests the possibility of constructing a machine for the development of useful electromotive forces. The essentials of a simple electromagnetic generator, or dynamo, are shown in Fig. 411.

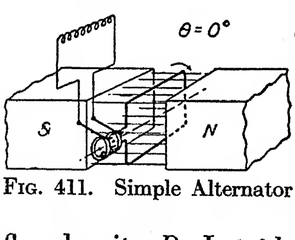


FIG. 411. Simple Alternator

The rectangular coil rotates about a horizontal axis with constant angular velocity ω in a uniform magnetic field of flux density B. Let θ be the angular displacement of the coil from the position in which it embraces maximum flux.

$$\omega \equiv \frac{\theta}{t}$$

$$\theta = \omega t. \qquad \text{(a)}$$

The short sides of the rectangle move always in planes parallel to the lines of flux, and hence do not cut across any lines. We need consider, therefore, only the motion of the long sides (i.e., those parallel to the axis of rotation), which are called **inductors** because emf's are induced in them.

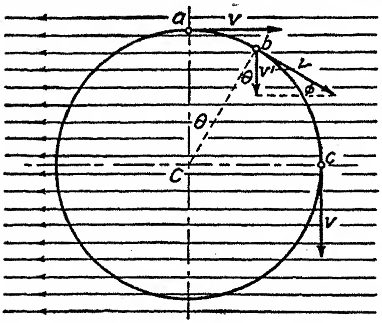


FIG. 412

In Fig. 412, the small circles represent the cross sections of inductors revolving about an axis perpendicular to the paper at C. For a general position such as b, the velocity v of the inductor is oblique to the flux, and cutting is due to the perpendicular component v' $(= v \sin \theta)$ only.

If l cm is the length of an inductor, the area swept over per second normal to the flux is

$$\frac{dA}{dt} = lv' \text{ cm}^2/\text{sec}$$

and the instantaneous time rate of cutting flux is

$$\frac{d\Phi}{dt} = B\frac{dA}{dt} = Blv' = Blv \sin \theta. \tag{b}$$

Since the rotating coil of Fig. 411 has two inductors in series, the instantaneous emf induced in the coil is, by the law of Henry and Faraday,

$$e = -N\frac{d\Phi}{dt} = -2Blv \sin \theta. \tag{c}$$

Here e has its maximum value $(\equiv E)$ when $\theta = 90°$ or $270°$, so that

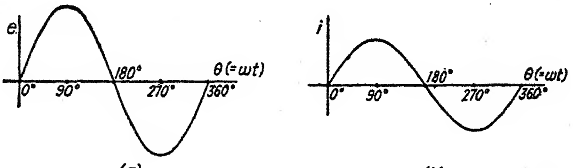
$$e_{90} \equiv E = -2Blv. \tag{d}$$

From Eqs. (a), (c), and (d),

$$e = E \sin \theta = E \sin \omega t \quad \text{abvolts.} \tag{381}$$

That is, the value of the emf at any instant (**instantaneous value**) is a sine function of the angle ωt, and has the maximum value E.

The plot of Eq. (381) is the "sinusoid," familiar from trigonometry. The values of e for one complete turn of the coil are called a cycle of values or, more briefly, **a cycle.** Fig. 413a shows a cycle for the above simple alternator. If the rotating coil is



Emf-Time Curve

Current-Time Curve

FIG. 413

connected by "slip rings" and "brushes," as shown, the instantaneous values of the current i in the circuit plotted against θ, or against t, will give a sine curve also (Fig. 413b).

By applying Fleming's generator rule to the inductors of the rotating coil, it will be seen that as the coil turns through 180°, from $\theta = 90°$ to $\theta = 270°$ (Fig. 414), the current changes sense in both the coil and the external circuit. The machine is therefore called an alternating current generator, or **alternator.**

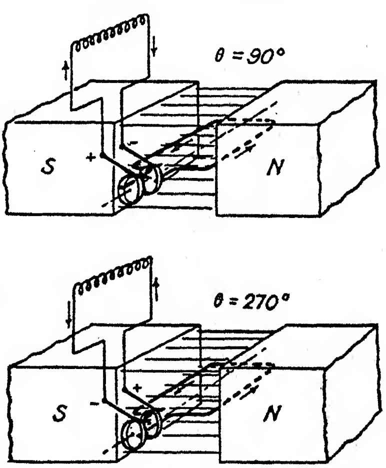


FIG. 414. In the Rotating Coil the Sense of the Current Reverses

Actual alternators have many inductors and fields produced by powerful electromagnets, but the principle of their operation is the same as that indicated above.

In the smaller sizes, the inductors are embedded in slots in a laminated iron core (the **rotor**) which rotates in front of the magnets that are fastened to the stationary frame (the **stator**). In the great alternators of industry, however, this arrangement is reversed: the inductors are mounted in the stator and the electromagnets are on the rotor. It is cheaper to build them in this way, and they are more satisfactory in operation.

Commercial alternators are generally designed with such a number of poles and speed of rotation that the emf will pass through a cycle of values (i.e., an inductor will pass by one N-pole and one S-pole) in 1/60 of a second. There are then 60 cycles of values in 1 sec, and hence the current is called **60-cycle current.**

There are various kinds of alternators:

1. *Single-phase* alternators have the inductors connected in a single continuous winding and brought out to two collector rings or terminals. They give emf-time curves like those in Fig. 413.

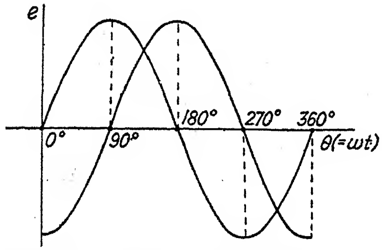


FIG. 415. Emf-Time Curve of Two Phase Alternator

2. *Two-phase* alternators have two separate single-phase windings brought out to two separate pairs of collector rings or terminals. The inductors of one winding are placed midway between those of the other winding; consequently the emf of one winding is "90° out of phase" with that of the other winding (Fig. 415).

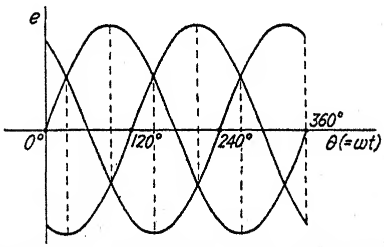


FIG. 416. Emf-Time Curve of Three Phase Alternator

3. *Three-phase* alternators have three single-phase windings which may be brought out to three separate pairs of rings or terminals, or may be connected so that only three rings or terminals are required. The inductors of any two phases divide the distance between successive inductors of the other phase into thirds. Hence the emf's of the three windings are "120° out of phase" with one another (Fig. 416).

457. Rectifiers. From the preceding section, it is evident that in electromagnetic generators the current produced is **always primarily alternating** in the coils of the armature.* In order to secure direct current, some form of *rectifier* must be used. The most generally used rectifiers are:

* Except in the case of the unipolar generator, which has never been a practical success.

1. The commutator (see Sec. 458).
2. The mercury-arc rectifier (Sec. 496).
3. The hot cathode rectifier (Sec. 512).
4. The copper-oxide rectifier (Sec. 497).

458. Direct-current generators. The alternating current developed in the armature of an electromagnetic generator may be delivered to the external circuit always in one direction (i.e., as direct current) by means of a commutator mounted on the armature shaft. The machine is then called a **direct-current generator.**

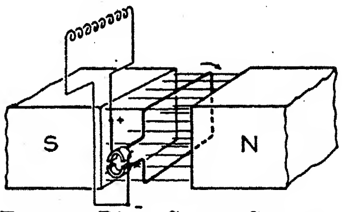


Fig. 417. Direct Current Generator

The **commutator** replaces the slip rings of the alternator. For a single coil it consists of a single ring split into halves (Fig. 417). The ends of the coil are soldered to these segments as shown. On tracing out the current in the circuit for a complete revolution of the coil, it will be seen that *as the coil passes through the position for which its emf is zero*, each brush passes from one segment to the other. That is, the connections of the rotating coil to the external circuit are reversed just when the direction of current in that coil reverses, and consequently the direction of the current in the external circuit remains unchanged.

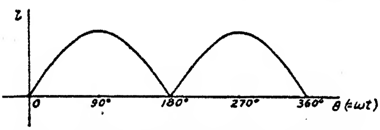


Fig. 418. Current-Time Curve for D. C. Generator of One Coil

The corresponding current-time curve for the direct current generator of one coil is shown in Fig. 418.

Commercial machines have many coils on the armature and a corresponding number of pairs of commutator bars (or segments). Each coil delivers energy to the external circuit except during the time that its commutator bars are under the brushes. The result is that the current in the external circuit is very nearly constant, there being only a slight *ripple* due to the finite width of the commutator bars (Fig. 419).

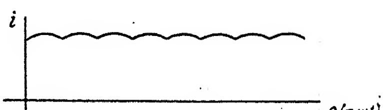


Fig. 419. Current-Time Curve for D. C. Generator of Many Coils

459. Generators as motors. Any direct-current generator will run as a direct-current motor, and vice versa; but differences in the design of motors and generators and in the adjustment of brushes may cause the efficiency to be less when either is used for the purpose for which it was not intended.

An alternating-current generator will run as a motor, when connected to an a-c source, if it is brought up to such a speed (**synchronous speed**) that the current in each conductor as it passes a pole is in the proper direction so that the torque on it will tend to rotate the armature always in the same direction. This type is called a **synchronous motor** and runs at a perfectly definite speed which depends upon the frequency of the impressed emf. Electric clocks that operate on ordinary 60-cycle lighting circuits are small synchronous motors.

460. Field excitation. The field magnets of an alternator are usually energized either from a small separate direct current generator driven by a belt from the shaft of the alternator (**separate excitation**), or by having a commutator in addition to its collector rings and rectifying part of its own current (**self-excitation**).

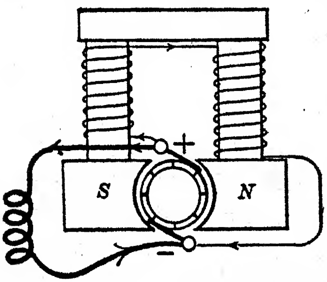


FIG. 420a. Shunt Generator. (Courtesy American Book Company)

The field magnets of a direct-current generator may be connected with its armature in four different ways:

1. *Shunt connection.* The field coils consist of many turns of fine wire and are connected as a shunt to the armature (Fig. 420a).

2. *Series connection.* The field coils consist of a few turns of large wire and are connected in series with the armature (Fig. 420b).

3. *Compound connection.* The field magnets have both shunt and series coils as above, the two being connected so as to produce the same polarity, i.e., to assist each other (Fig. 420c).

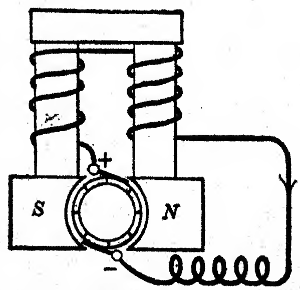


FIG. 420b. Series Generator. (Courtesy American Book Company)

4. *Differential connection.* This is the same as compound, but with the shunt and series coils opposing each other.

The characteristics of generators connected in these four ways are discussed in any text on dynamoelectric machines. The compound connection is the one most generally used.

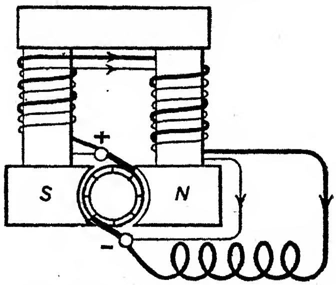


FIG. 420c. Compound Generator. (Courtesy American Book Company)

461. Back emf in motors. When a machine is running as a motor, its conductors are cutting across the lines of flux of the field just as if the armature were being rotated by mechanical means. Hence there is generator action in every running motor; and the emf E_b generated is called the *back emf*, since by Lenz's law it opposes the cause, which is the impressed emf.

This may be shown by connecting an electric lamp across the terminals of a motor having a heavy flywheel. The lamp will continue to glow after the motor has been disconnected from its source of energy, for the inertia of the flywheel will keep the armature turning, and the back emf of the motor will produce sufficient current to maintain the incandescence of the filament until the energy of the flywheel is dissipated. A voltmeter used in place of the lamp will show that there is no change of polarity when the motor is disconnected from the mains. Therefore, the back emf must have been in opposition to the impressed emf.

462. Transformation of energy in d-c motors. Consider a d-c generator of emf E, which produces a current I in a circuit containing a *series* motor whose back emf is E_b and resistance R. Then the

$$\text{Power by the impressed emf from generator} = EI$$
$$\text{Power by back emf of motor} = E_bI$$
$$\text{Net potential difference causing current in motor} = E - E_b$$
$$\text{Current through motor } I = \frac{E - E_b}{R}$$

Energy lost as heat in motor is

$$RI^2t = (E - E_b)It = EIt - E_bIt.$$

Since the RI^2t loss represents practically all the electrical loss in a motor,

$$\text{Electric energy transformed into mechanical energy is } EIt - RI^2t = E_bIt. \tag{382}$$

That is, **the electrical energy that is transformed into mechanical energy in a series motor equals the product of the back emf of the motor, the current through the armature, and the time.** This is true of any d-c motor, though the proof for other types is more involved. It is not true for a-c motors on account of the effects of capacitance and inductance.

PROBLEMS

1. A generator develops 20 amp at 100 volts. What is the power in kilowatts and what is the horsepower?

2. The armature of a certain generator has 200 conductors in series between the brushes. Each conductor cuts 3,000,000 lines of flux 1500 times per minute. What is the average emf developed?

3. An electric motor requires 12 amp at 110 volts. The two wires leading from the generator to the motor have a resistance of 0.2 ohm each. What terminal potential must the generator supply? How much power is lost in the line?

4. An electric motor requires 8 amp at 110 volts. The two wires leading from the generator to the motor have a resistance of 0.3 ohm each. What terminal potential must the generator supply? How much power is lost in the line?

5. A generator is to develop a current of 25 amp at 110 volts. If its efficiency is 80%, what must be the horsepower of an engine to drive it? What will be the torque on its shaft when running at 1800 rpm?

6. The winding of the armature of a motor has a resistance of 0.5 ohm and may have a maximum safe current of 10 amp. How much resistance must be in the starting box to protect the armature at starting, when connected to 110-volt mains? When this resistance has been cut out, what is the back emf?

7. A 2-hp motor, when connected across 220-volt mains, requires 8 amp. What is its efficiency?

8. If the speed of the above motor is 3600 rpm, what torque does it develop?

√**9.** If the resistance of the armature of a shunt motor is 0.4 ohm and the current in it is 25 amp when it is connected across 110-volt mains, what is the back emf of the motor and its theoretical horsepower?

√**10.** A certain series motor has an armature resistance of 0.6 ohm and field coils whose resistance is 5 ohms. If the current is 9 amp when the motor is connected across 110-volt mains, what is the back emf?

√**11.** In problem 10, what is the input power, the output power, and the power wasted in heat?

√**12.** A shunt motor has an armature resistance of 0.5 ohm and a field resistance

of 60 ohms. When first connected to 30-volt mains, what is the current in the armature, in the field coils, and the total current?

13. If the above motor takes 5 amp when running at full speed, what is the back emf? The useful power? The power wasted in heat?

14. The resistance of the armature of a certain shunt motor is 0.2 ohm and that of the field, 150 ohms. If the safe armature current is 20 amp, what resistance must be connected in series with the armature at starting when connected to 110-volt mains? If the current in the armature at full speed is 8 amp, what is then the back emf?

CHAPTER XXXIII

INDUCTANCE

463. Self-inductance. If a current i of electricity is established in a coil, it will produce a magnetic flux through the coil. Let the coil be close wound (Fig. 421) so that substantially all the lines of flux Φ pass through all N turns. We should expect that if the current were doubled, the flux would be doubled; if the current were trebled, the flux would be trebled; and so on, as shown in the table below.

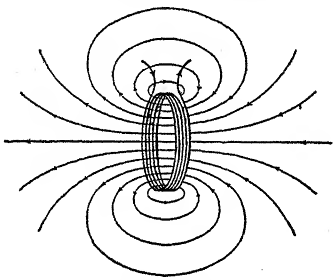


FIG. 421

Current	Flux
i	Φ
$2i$	2Φ
$3i$	3Φ
⋮	⋮

Experiment shows this proportionality to be true, **provided there is no iron or other magnetic material in the core of the coil.**

Hence we may write:

$$\Phi \propto i \tag{a}$$

and therefore

$$N\Phi \propto i$$

whence

$$N\Phi = Li \tag{b}$$

where the factor of proportionality L is called the *self-inductance*, or merely the *inductance*, of the coil.

Solving Eq. (b) for L, we have the definition:

$$L \equiv \frac{N\Phi}{i}. \tag{383}$$

The product $N\Phi$ is called the number of **flux-linkages,** or flux-turns, associated with the coil, a flux-linkage consisting of a line

of flux and a turn of wire each passing through the other like the links of a chain. In words, therefore, **the self-inductance of a coil is the change of flux-linkages due to a unit change of current in the coil itself.**

If the coil contains iron or other ferromagnetic material, the proportionality of Eq. (a), above, is not true because the permeabilities of such materials are not constant. But if we take sufficiently small steps, the change of flux $d\Phi$ is proportional to the corresponding change of current di, so that we may write:

$$d\Phi \propto di$$

$$Nd\Phi \propto di$$

$$Nd\Phi = Ldi \tag{c}$$

$$L \equiv \frac{Nd\Phi}{di}. \tag{384}$$

It will be seen from Fig. 422 that in both the above definitions L is the product of the number of turns N by the slope of the Φ-i curve. Hence Eq. (384) includes Eq. (383) as a special case in which the slope and therefore the inductance is constant. But in the case of ferromagnetic materials, the permeability is not constant, hence **the slope and L are not constant.** In practical problems, we assume the permeability to be constant at its average value and compute the corresponding value of L.

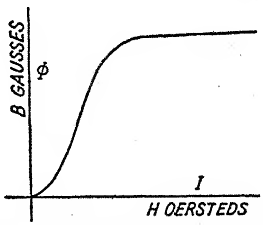


FIG. 422. Φ-I Curve for Iron

Since, when i changes in the wire, the flux Φ through the coil changes, there will be an emf induced in the coil. This **self-induced emf** may be found as follows:

By the law of Henry and Faraday, Eq. (308),

$$e = -N\frac{d\Phi}{dt}.$$

From Eq. (c), above,

$$Nd\Phi = Ldi.$$

Therefore,

$$e = -L\frac{di}{dt} \tag{385}$$

where e is the emf developed in the coil when the current changes in the coil itself. The $(-)$ indicates that this emf opposes the change of current; hence it is called the **back emf of self-induction.**

In Eq. (370) it is seen that for a change of current di in a coil, the change of flux $d\Phi$ is itself proportional to N. Hence self-inductance is proportional to N^2: if we double the number of turns, other things being constant, we quadruple the inductance.

Most coils are not close wound, and consequently the total flux Φ does not interlink with all N turns (Fig. 423). In such coils, $N\Phi$ of the preceding discussion becomes the summation of the flux-linkages of the individual turns $\Sigma N\Phi$. Hence $\Sigma N\Phi$ and the self-inductance L will depend upon the shape and size of the coil, the way it is wound, and the permeability of the core.

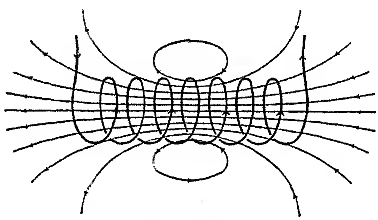


FIG. 423

Except in a few special cases, the calculation of inductance is quite an involved process.* Inductance is readily measured, however, by means of an inductance bridge or by a wave-meter and oscillator.†

Mechanical analogue of self-inductance. When a circuit is closed and the current begins to increase, the emf due to self-inductance opposes the impressed emf (from the battery or other source of supply) and hence slows up the change of current. Similarly, when the impressed emf is reduced and the current tends to decrease, the self-induced emf opposes the change and tends to maintain the current.

The effect of self-inductance L on the motion of electricity is therefore quite analogous to that of mass in translation and to moment of inertia in rotation; or we may say that **self-inductance in electricity is analogous to inertia in mechanics.** Researches of recent years into the electrical nature of matter make it appear probable that all inertia effects are, in the last analysis, due to the self-inductance of electric charges.

Self-inductance is shown by the following experiment. Let an

* See *Bulletin* of U. S. Bureau of Standards, Vol. VIII (1912), No. 1.

† Terry and Wahlin, *Advanced Laboratory Practice in Electricity and Magnetism* (New York, McGraw-Hill Book Co., 1936), p. 120.

ordinary 6-volt automobile lamp and a coil of low resistance but large inductance (i.e., having many turns on an iron core) be connected in parallel across a 6-volt battery, with a variable resistance r in series, as shown in Fig. 424.

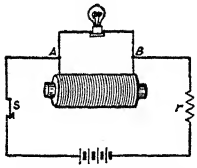


FIG. 424. Self-Induced Emf in the Coil Lights the Lamp

When the switch is closed we should expect the lamp to glow momentarily, but usually this will not be noticeable because of the time required for the filament to heat up. As soon as the current reaches a steady value, the back emf of self-induction disappears. The resistance of the coil being small, the ri-drop between A and B is then only sufficient to make the lamp glow a dull red.

When the circuit is broken at S, however, a large self-induced emf is developed in the coil, which tends to maintain the current and is sufficient to make the lamp flash up with full brightness.

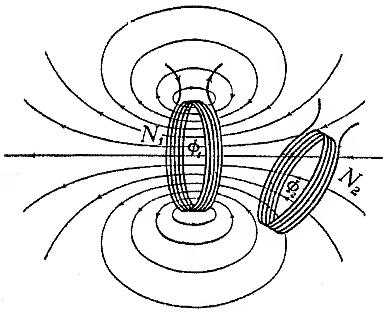


FIG. 425. Mutual Inductance

464. Mutual inductance. Consider two coils (Fig. 425) having N_1 and N_2 turns, respectively, and so placed that a part of the flux Φ_1 due to the current i_1 in coil N_1 interlinks with all the turns of coil N_2. Let this flux through the coil N_2 be Φ_2.

Then, if i_1 changes, Φ_1 and consequently Φ_2 will change. If i_1 is doubled, we might expect Φ_2 to be doubled; if i_1 is trebled, Φ_2 to be trebled; and so on as in the table below.

Current in n_1	Flux in n_2
i_1	Φ_2
$2i_1$	$2\Phi_2$
$3i_2$	$3\Phi_2$
⋮	⋮

This is found by experiment to be correct, **provided there is no iron or other ferromagnetic material in the coils**; i.e., if the permeability is constant.

Hence we may write:

$$\Phi_2 \propto i_1 \tag{a}$$

$$N_2\Phi_2 \propto i_1$$

$$N_2\Phi_2 = M_{12}i_1 \tag{b}$$

and the factor of proportionality M_{12} is called the *mutual inductance* of coil N_1 with respect to coil N_2.*

From Eq. (b) we get the definition of mutual inductance:

$$M_{12} \equiv \frac{N_2\Phi_2}{i_1}. \tag{c}$$

Expressed in words, this is: The mutual inductance M_{12}* of coil N_1 with respect to coil N_2 is the change of flux-linkages produced in coil N_2 by a unit change of current in coil N_1.

If the coils contain iron or other ferromagnetic materials, the proportionality of Eq. (a), above, is not true because their permeabilities are not constant (Fig. 422), and the definition must be written:

$$M_{12} \equiv \frac{N_2 d\Phi_2}{di_1}. \tag{386}$$

Reversing the roles of the coils, the mutual inductance of coil N_2 with respect to coil N_1 is:

$$M_{21} \equiv \frac{N_1 d\Phi_1}{di_2}. \tag{387}$$

The permeability μ of ferromagnetic material is a complicated function of the magnetizing current, so that for coils containing such material M_{12} and M_{21} are not constant nor always equal.

If the permeability may be considered constant, as is the case in most practical work, it may be shown that

$$M_{12} = M_{21} \equiv M \tag{d}$$

and M is called the **mutual inductance of the two coils.**

The change of flux in either coil induces an emf in the other coil which opposes the change of current that produces it. Its

* Some writers call this the mutual inductance of coil N_2 with respect to coil N_1 and write the symbol: M_{21}.

value may be found by the law of Henry and Faraday, as in the case of self-inductance.

By Eq. (308),

$$e_2 = -N_2 \frac{d\Phi_2}{dt}.$$

From Eq. (386),

$$N_2 d\Phi_2 = M_{12} di_1$$

which, substituted above, gives:

$$e_2 = -M_{12}\frac{di_1}{dt} = -M\frac{di_1}{dt}. \qquad (388)$$

And when coil N_2 is the primary and coil N_1 the secondary, we have the similar equation:

$$e_1 = -M_{21}\frac{di_2}{dt} = -M\frac{di_2}{dt} \qquad (389)$$

where e_2 and e_1 are the emfs induced in the respective coils by the change of current in the other coil.

465. Units of inductance. The units of inductance are usually defined from Eqs. (385) and (388). From the former,

$$e = -L\frac{di}{dt}$$

where L will be unity when e, di, and dt are all unity; that is,

$$1 \text{ abvolt} = -1 \text{ abhenry} \times \frac{1 \text{ abampere}}{1 \text{ second}}. \qquad \text{(a)}$$

From Eq. (388),

$$e_2 = -M\frac{di_1}{dt}$$

where M will be unity when e_2, di_1, and dt are all unity; that is,

$$1 \text{ abvolt} = -1 \text{ abhenry} \times \frac{1 \text{ abampere}}{1 \text{ second}}. \qquad \text{(b)}$$

From Eqs. (a) and (b) it is seen that the unit of L is the same as the unit of M; and in the electromagnetic system it is called

the *abhenry*. Describing in words this unit as defined by these equations, we say, from Eq. (a):

The **abhenry** is the **self**-inductance of a coil in which a back emf of one abvolt is induced when the current in that coil changes at the rate of one abampere per second. And from Eq. (b):

The **abhenry** is the **mutual** inductance of two coils if an emf of one abvolt is induced in one coil when the current changes at the rate of one abampere per second in the other.

Similarly, by using practical units throughout Eqs. (a) and (b), we get the definitions of the practical unit of inductance, the *henry:*

The **henry** is the **self**-inductance of a coil in which a back emf of one volt is induced when the current in that coil changes at the rate of one ampere per second.

The **henry** is the **mutual** inductance of two coils if an emf of one volt is induced in one of them when the current changes at the rate of one ampere per second in the other.

The relations between the electromagnetic and the practical units of inductance are easily found as follows.

From Eq. (a), above,

$$1 \text{ abhenry} \equiv \frac{1 \text{ abvolt}}{1 \text{ abamp}/1 \text{ sec}} \tag{c}$$

and similarly,

$$1 \text{ henry} \equiv \frac{1 \text{ volt}}{1 \text{ amp}/1 \text{ sec}}. \tag{d}$$

But

$$1 \text{ volt} \equiv 10^8 \text{ abvolts}$$

and

$$1 \text{ ampere} \equiv \tfrac{1}{10} \text{ abampere}.$$

Substituting these values in Eq. (d),

$$1 \text{ henry} = \frac{10^8 \text{ abvolts}}{\frac{1}{10} \text{ abamp}/1 \text{ sec}}$$

$$= \frac{10^9 \text{ abvolts}}{1 \text{ abamp}/1 \text{ sec}}$$

which, by comparison with Eq. (c), above, gives:

$$1 \text{ henry} = 10^9 \text{ abhenries}. \tag{390}$$

As will be seen later, inductance, like resistance, tends to impede the flow of electrons in alternating currents, though in a different way. Hence it is a help to the memory to note here that **the relation between the corresponding resistance units is the same as for inductance units:** viz., by Eq. (315),

$$1 \text{ ohm} = 10^9 \text{ abohms.}$$

466. Discussion of inductance. From the foregoing sections, it is clear that the emf's due to self- and mutual inductance are proportional to the **rate of change** of current. Hence these **inductive effects occur only when the current is changing.**

Since direct currents are generally used at steady values, **inductance is negligible in direct-current work, except when the current is turned on or off.** That is why we always close the battery key of a Wheatstone bridge first. This gives the current time to grow to its steady value before the galvanometer key is closed.

If we close the galvanometer key and then the battery key, there will be a back emf due to the inductance of the coils as the current grows in the circuit. This will cause the galvanometer to give a kick, even though the bridge is actually in balance. A similar kick in the opposite direction will be obtained if the battery key is opened before the galvanometer key, even though the bridge is balanced.

With alternating currents, however, the current is changing continually except at the instants when it has its maximum and minimum values. Consequently, **inductance plays a very important role in alternating currents,** its effect in some instances completely overshadowing the effect of resistance. Since the instantaneous emf due to inductance always opposes the change of current, this emf will sometimes oppose, sometimes assist, and sometimes not affect the impressed emf. Just how these effects occur will depend upon the phase difference of the two emf's, which will be taken up later. It may be said, however, that the net emf is the *vector sum* of the impressed emf and the emf due to inductance.

Numerical values of inductance are not so easily calculated as are those for resistance. The reason is that inductance, as has been seen, depends not only upon the number of turns but also

upon the shape of the coil, how it is wound, and the kind of core within it; while, in addition, mutual inductance depends upon the relative positions of the coils. It is therefore not feasible to consider here the inductance of coils in series and in parallel.*

467. Energy of a magnetic field. In mechanics (Sec. 79), we saw that work done in overcoming inertia is stored in the system as kinetic energy and has the values:

$$\text{K.E.} = \tfrac{1}{2}Mv^2 \quad \text{for translation}$$

and

$$= \tfrac{1}{2}I'\omega^2 \quad \text{for rotation.}$$

Since inductance has appeared to be the electrical analogue of inertia, we might expect that energy would be stored up in a magnetic field, because in building up a current in a coil, work must be done against the back emf of inductance as the flux through the coil increases. It was this stored energy that caused the lamp in Fig. 423 to glow brightly when the circuit was broken.

The energy stored in a magnetic field is equal to the work done in building up the field; i.e., in building up the current with which the field is associated. At any instant during the process, let the value of the growing current be i. When this changes by an amount di, there develops a back emf $\left(e_b = -L\frac{di}{dt}\right)$, which opposes the change.

In order to produce the change of current, we must therefore impress upon the circuit an emf e equal and opposite to e_b. The work done upon the system in the time dt will then be:

$$dw = eidt = (-e_b)idt = \left(L\frac{di}{dt}\right)idt = Lidi$$

and the total work to build up the current from 0 to its final value I will be:

$$W = \int_0^I Lidi = \left[L\frac{i^2}{2}\right]_0^I = \frac{1}{2}LI^2 \qquad (391)$$

* For discussions of these questions, the student is referred to the following:
Bulletin, U. S. Bureau of Standards, *op. cit.*
Page and Adams, *Principles of Electricity* (New York, D. Van Nostrand, 1931).
Morecroft, *Principles of Radio Communication* (New York, John Wiley & Sons, 1927).

which is correct for either emu or practical units, provided the same system is used throughout.

Equation (391) may be looked upon as an expression for electrical kinetic energy. It is seen to have the same form as the above equations for mechanical kinetic energy, self-inductance corresponding to mass M in translation and to moment of inertia I' in rotation.

CHAPTER XXXIV

ALTERNATING CURRENTS AND INSTRUMENTS

468. Advantages of using alternating currents. In Sec. 419, it was shown that when electricity is forced through a circuit against resistance, the energy lost as heat is:

$$W_1 = RI^2t.$$

From this it is seen that, in a given resistance, the smaller the current the less energy is wasted.

If the emf impressed on the circuit is E, the total energy W supplied is:

$$W = EIT$$

so that for a given W, if I is decreased, E must be correspondingly increased.

The following solved problem may make the matter clearer.

Solved Problems

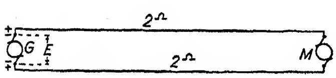


FIG. 426. Transmission Line

A transmission line (Fig. 426), whose total resistance is 4 ohms, receives energy at the rate of 1 kw from a generator G and delivers energy to a motor M. What will be the efficiency of transmission if the output voltage of the generator is (a) 100 volts? (b) 1000 volts?

CASE I. *Given:* Output voltage of generator = 100 volts

Output power of generator = 1000 watts.

Solution: By Eq. (336),

$$P = EI.$$

Therefore,

$$1000 = 100I$$

$$I = 10 \text{ amp} = \text{current produced by generator.}$$

By Eq. (338), power lost in overcoming line resistance is

$$\begin{aligned} P &= RI^2 \\ &= 4 \times (10)^2 \\ &= 400 \text{ watts.} \end{aligned}$$

$$\text{Efficiency of transmission} = \frac{1000 - 400}{1000} = 60\%.$$

CASE II. *Given:* Output voltage of generator = 1000 volts
Output power of generator = 1000 watts.

Solution: Current produced by generator $= \dfrac{1000 \text{ watts}}{1000 \text{ volts}} = 1$ amp

Power lost in overcoming resistance $= 4 \times (1)^2 = 4$ watts

$$\text{Efficiency of transmission} = \frac{1000 - 4}{1000} = 99.6\%.$$

Hence, for the most economical transmission of energy, low currents and high voltages are used.

The fact that, by means of the alternating-current transformer, power may be changed from low voltage and high current to high voltage and low current (and vice versa), with great ease and efficiency, has caused alternating current to take the place of direct current for the great majority of purposes, particularly for transmission over long lines. A further advantage in its use is that alternating current generators and motors can be built more economically and in general give more satisfactory service than do direct-current motors.

Direct current must obviously be used for electrolytic purposes; and as yet direct-current motors are superior to alternating-current motors where variable speed is required.

469. Effective value of alternating current. Twice during a cycle, the value of an alternating current varies from zero to its maximum value I and back to zero (Fig. 427). The question naturally arises: What shall we consider the effective value of such a current?

Most electrical problems are concerned primarily with getting work done. It is therefore logical to say that an alternating current has an effective value of 10 amp when it yields as much work as does a direct current of 10 amp in the same length of time.

If an emf E is employed only in maintaining a current I in a

resistance R, all the energy is converted into heat in accordance with Joule's law:

$$H = .24RI^2t.$$

We therefore say that an alternating current has a value of 10 amperes when it will produce in a given resistance as much heat per second as is produced per second in that resistance by a direct current of 10 amperes.

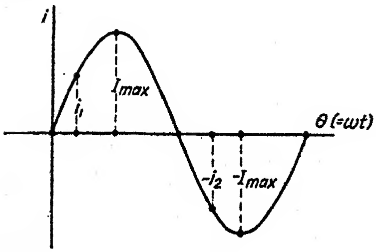


FIG. 427

When the instantaneous value of the alternating current is i_1, it produces heat proportional to i_1^2; and when its value is $-i_2$, it produces heat proportional to i_2^2; and so on for all the values in the cycle. Hence, for the whole cycle, the **effective value of the current** for producing heat, i.e., for transmitting energy, is:

$$I_{\text{eff}} = \sqrt{\text{Average } i^2}.$$

By means of calculus it is easily shown * that

$$I_{\text{eff}} = \tfrac{1}{2}\sqrt{2}I_{\max} = 0.707I_{\max}. \tag{392}$$

Multiplying through by the resistance R of the circuit, we have:

$$RI_{\text{eff}} = 0.707RI_{\max}$$

or

$$E_{\text{eff}} = 0.707E_{\max}. \tag{393}$$

Now it so happens that when we make alternating-current instruments in the easiest and most obvious ways, they invariably indicate effective values of current or emf, as defined above. Hence in general, when an alternating current or emf is stated, **effective value** is meant. Other names for effective value are "virtual value" and "root-mean-square (rms) value."

470. The electrodynamometer. If a direct-current ammeter or voltmeter having a moving coil and a **permanent** magnet (see Fig. 351) were connected to an a-c source, the pointer would merely tremble. This is explained by the fact that the polarity

* Christie, *Electrical Engineering* (New York, McGraw-Hill Book Co., 1917), p. 106.

of the moving coil would reverse each time that the current reverses, hence the torque on the coil would likewise reverse, and the inertia of the coil would prevent it from making more than a trivial deflection before the next reversal.

The obvious way to overcome this difficulty is to produce the stationary magnetic field by coils of wire instead of by a permanent magnet. If these coils are energized by some of the same alternating current that energizes the moving coil, their polarities and that of the moving coil will reverse at the same time, and hence the torque will be always in one direction. When this is done, the type of galvanometer so obtained is called an *electrodynamometer* (Fig. 428). Unfortunately, the deflection is proportional to the square of the effective current and not to the effective current. Hence, when the scale is divided to indicate equal increments of current, the angular divisions are not equal; i.e., we do not get a uniformly divided scale.

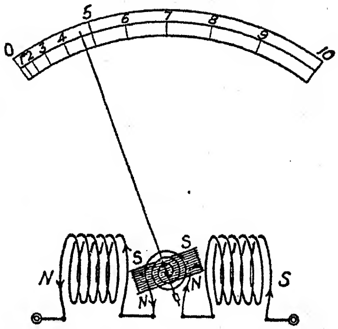


FIG. 428. Electrodynamometer Galvanometer

Electrodynamometer instruments will obviously measure direct as well as alternating current. It may be made into an ammeter by connecting a shunt across its terminals, or into a voltmeter by placing a multiplier in series with its coils, as in Sec. 398. Since there is no ferromagnetic material in their construction, the calibration of such instruments is the same for a-c as for d-c. **They may be calibrated on a d-c potentiometer and then used as secondary standards for the calibration of other a-c instruments.** They are not suitable for currents and emf's of high frequency, as their inductance then introduces a considerable error.

471. Soft-iron instruments. Probably the most generally used type of alternating current ammeter and voltmeter is the so-called soft-iron, or iron-vane, type. These instruments make use of the fact that soft iron has very low retentivity; i.e., it loses its magnetism very promptly when the magnetizing force is removed.

A small segment of a cylinder of soft iron J (Fig. 429) is mounted on a staff with the pointer P, and these constitute the moving

system. This moving system turns about the axis of a coil of wire C in which it is mounted. To the inner surface of this coil is fastened another piece of soft iron K.

When the coil is energized with either d-c or a-c, both the movable piece and the fixed piece of soft iron are magnetized, with their top ends of the same polarity and their bottom ends of the same polarity. These poles repel; and since the only motion possible is rotation of the movable vane, this system moves into a position where the deflecting torque due to the repelling poles is counterbalanced by the restoring torque due to a spiral spring.

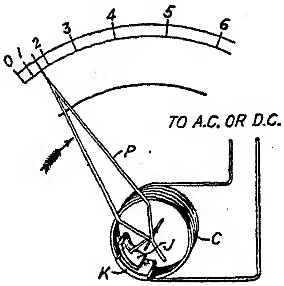


FIG. 429. Soft Iron Galvanometer. (Weston Electrical Instrument Co.)

The deflecting torque is proportional to the product of the pole strengths of the two pieces of soft iron, each of which is proportional to the current in the coil. The deflection is therefore proportional to the average squared current, as in the case of the electrodynamometer; and hence **the instrument will indicate effective current.** But the scale will not be equally divided, and it should be noted that the calibration is not quite the same for a-c and for d-c on account of inductance. For the same reason, these instruments are not suitable for high-frequency currents.

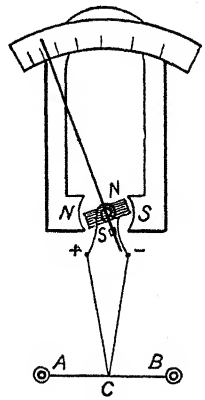


FIG. 430. Thermoelectric Instrument

472. Thermoelectric instruments. These instruments consist essentially of a piece of thin resistance wire AB (Fig. 430) a few centimeters long, which is heated by the whole or a part of the current to be measured. At the center of this wire is soldered the hot junction of a thermocouple. The emf developed by the thermocouple is a function of the temperature of the hot-wire, and is measured by a millivoltmeter of the permanent magnet type (Sec. 398), which forms an integral part of the instrument. The sensitivity is greatly increased if the hot-wire is enclosed in a glass tube exhausted to a high vacuum.

From the definition of effective value, an alternating current having an effective value of, say, 10 amp will obviously heat a wire as much as will a direct current of 10 amp in the same length of time. These instruments may therefore be calibrated to read effective values, and the calibration is the same for a-c as for d-c.

They are satisfactory for currents of either **high or low frequency,** since the inductance, which is only that of a straight wire, is very small. For this reason, they are the only type of instrument suitable for currents of radio frequency.

Thermoammeters are made by using a shunt in parallel with the hot-wire; and thermovoltmeters, by a multiplier in series with it, as in the case of other types.

473. Emf-current relation for alternating currents. Consider a coil of large wire (say, No. 12) having several hundred turns so wound as to leave through its center a hole 5 or 6 cm in diameter. Let its resistance, as determined with a Wheatstone bridge (i.e., with d-c), be 2 ohms.

If this coil is connected to 110-volt d-c mains with an ammeter in series, we should expect the current to be 55 amp, in accordance with Ohm's law; and so it is.

But if the same coil is connected to 110-volt a-c mains, with an a-c ammeter in series, instead of 55 amp we will read perhaps 7 amp. Moreover, as laminations (thin sheets) of wrought iron are placed through the hole in the coil, the current will decrease until hardly readable. Clearly **Ohm's law does not tell the whole truth in the case of alternating currents.** The addition of laminations caused the current to decrease, though obviously it did not change the resistance of the coil. **What it did do was to change the inductance.**

If a group of condensers is added in series with the above coil, their capacitance may be so adjusted as to bring the current back to its original value (7 amp), and theoretically even to the value 55 amp, as predicted by Ohm's law.

This experiment shows that inductance and capacitance, as well as resistance, play a part in determining the magnitude of the alternating current in a circuit. It may be shown * that for series connections the relation is:

* Christie, *op. cit.*, p. 114.

$$I = \frac{E}{\sqrt{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}} \tag{394}$$

where I is the effective current in amperes;
E is the effective emf in volts;
R is the pure resistance in ohms;
L is the inductance in henries;
C is the capacitance in farads; and
$\omega = 2\pi n$ where
n is the frequency of the alternating current.

By analogy with Ohm's law, Eq. (394) is often called **Ohm's law for alternating currents,** although it was formulated many years after Ohm's death.

The term ωL is called the *inductive reactance;*

$\dfrac{1}{\omega C}$ is called the *capacitative reactance;*

and the whole denominator, $\sqrt{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}$, is called the *impedance.*

It will be observed that Eq. (394) reduces to Ohm's law when the inductive and the capacitative reactances are both zero, or when they are equal.

474. The circle diagram. To avoid the labor of plotting sine-wave curves similar to those of Fig. 413 for each alternating emf and current, the following simpler device, known as the *circle diagram,* is employed.

A circle (Fig. 431) is drawn whose radius to a convenient scale represents the maximum value of the alternating emf under consideration. A radius vector of this circle is labeled E and is considered to rotate counterclockwise with a uniform angular speed ω equal to that which a simple alternator (Fig. 412) would require in order to develop an alternating emf of the same frequency n as that being studied. That is,

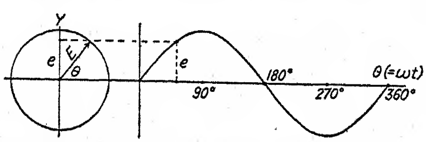


FIG. 431. The Circle Diagram

$$\omega = 2\pi n.$$

At any time t after passing the initial line, this radius vector will have turned through an angle θ ($= \omega t$), and its projection on the Y-axis will be:

$$E \sin \theta = E \sin \omega t.$$

But, by Eq. (381),

$$e = E \sin \omega t.$$

That is, the projection of the radius vector E on the Y-axis at any time t gives the value of the scalar *e at that instant.* A similar convention applies to the representation of instantaneous values of alternating currents. Radii vectores representing several emf's and their corresponding currents are often drawn in the same circle diagram.

The circle diagram has many properties not at first obvious, and it has become the basis of a powerful method of analysis which is employed in the majority of alternating-current problems.

475. Phase difference. From Sec. 463, it will be recalled that inductance plays the same role in electrical circuits that inertia plays in mechanics. In order to give shm to a large mass M on a frictionless, horizontal plane, we must push with maximum force F to the right when the body is at A (Fig. 432), whereas the maximum velocity v to the right does not occur until M is at C, which is 1/4 period, or 90°, later. That is, v lags 90° behind F.

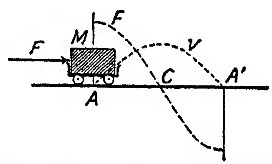


FIG. 432. Velocity Lags behind Force

Similarly, when a harmonic alternating current is maintained in a circuit having small resistance R and great inductance L, the current I lags 90° in phase behind the emf E.

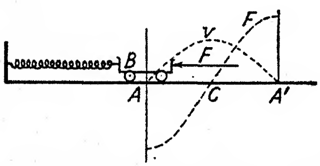


FIG. 433. Velocity Leads Force

Capacitance is the electrical analogue of elasticity. Let a body B (Fig. 433) of very small mass be attached to a light but stiff spring and rest on a frictionless, horizontal plane. If we give B shm, the maximum force F to the right must be exerted at A'; whereas the maximum velocity to the right v occurred at

C, which is 1/4 period, or 90° in phase, ahead of A'. That is, v leads F in phase by 90°.

Similarly in electricity, if a circuit has small resistance and large capacitance, the current will reach its maximum value 1/4 period before the emf reaches its maximum; that is, current I leads emf E in phase by 90°.

These analogies are confirmed by means of the **oscillograph.** This is a sensitive galvanometer of the D'Arsonval type, having a very short period and so little inertia that its motion follows accurately the changes of alternating current. Figures 434–437 show drawings of oscillograms for the various circumstances. Emf curves are shown by long dashes; current curves, by short dashes. The corresponding circle diagrams are shown at the left on each figure.

Fig. 434. The circuit has **resistance** only, the inductance and capacitance being negligible. I and E are in phase, i.e., $\phi = 0$.

Fig. 435. The circuit has large **inductance,** the resistance and capacitance being negligible. I lags 90° behind E, i.e., $\phi = 90°$.

Fig. 436. The circuit has large **capacitance,** resistance and inductance being negligible. I leads E by 90°, i.e., $\phi = -90°$.

Fig. 437. The circuit has **resistance, inductance,** and **capacitance** of such values that I lags 60° behind E, i.e., $\phi = 60°$.

The angle ϕ is called the **phase difference** between E and I. It is the phase angle ϕ_E of the emf minus the phase angle Φ_I of the current.

476. Alternating-current power. In Sec. 420 it was shown that with direct current the power is:

$$P = EI.$$

Obviously for alternating current the instantaneous power is

$$p = ei \tag{395}$$

where e and i are the values of the emf and current at the instant in question.

Since both e and i are sine functions of time, the instantaneous power p will vary throughout the cycle, whether e and i are in phase or not. The average value of the power throughout a cycle

must therefore be obtained by integration. By this method * it is found that, *for alternating currents*,

$$P = EI \cos \phi. \tag{396}$$

The results of this relation are rather surprising, as is illustrated by the accompanying figures.

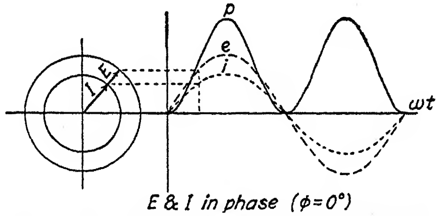


FIG. 434. Power Curve—E and I in Phase

Fig. 434.—When I and E are in phase, as they are when the circuit has resistance only or when the inductive reactance just offsets the capacitative reactance, $\phi = 0$; $\cos \phi = 1$; and the power is all positive. That is, the circuit is receiving energy from the mains throughout the entire cycle.

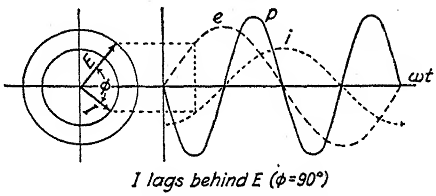


FIG. 435. Power Curve—I Lags 90° behind E

Fig. 435.—When I lags 90° behind E, as it does in a circuit having inductance only, $\phi = 90°$; $\cos \phi = 0$; and *the average power during a cycle is zero.* That is, the circuit receives energy from the mains during one half-cycle, and returns an equal amount to the mains during the other half-cycle.

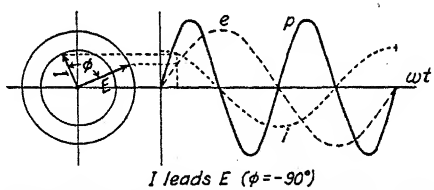


FIG. 436. Power Curve—I Leads E by 90°

Fig. 436.—When I leads E by 90°, as it does in a circuit having capacitance only, $\phi = -90°$; $\cos \phi = 0$; and *the average power during a cycle is again zero.* The results are the same as in the preceding case, with the half-cycles reversed.

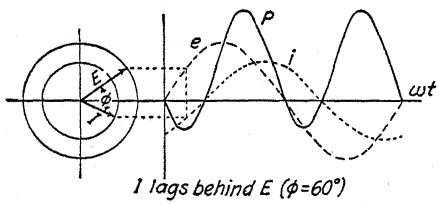


FIG. 437. Power Curve—I Lags 60° behind E

Fig. 437.—When I lags behind E by any angle, say, 60°, $\phi = 60°$; $\cos \phi = 1/2$; and $P = 1/2EI$. This repre-

* N. E. Gilbert, *Electricity and Magnetism* (New York, The Macmillan Co., 1932), p. 244.

sents a general case in which a circuit has resistance, inductance, and capacitance. The circuit receives energy during part of the cycle and returns energy to the mains during the other part of the cycle.

477. Power factor. In the case of most a-c circuits, we do not know at first the value of ϕ. If we multiply the reading E of an a-c voltmeter across the circuit by the reading I of an a-c ammeter in series in the circuit, we get a value EI, which is the *apparent power* P' in volt-amperes. If, for the same circuit, we measure the true power P by means of a wattmeter, we should have:

$$\frac{P}{P'} = \frac{EI \cos \phi}{EI} = \cos \phi$$

or

$$P = P' \cos \phi. \tag{397}$$

Hence $\cos \phi$ is called the **power factor** of the circuit, for it is the factor by which the apparent power in volt-amperes must be multiplied in order to find the true power in watts.

478. The wattmeter. The electrodynamometer (Sec. 470) may easily be converted into a wattmeter. For this purpose, the fixed coils are made of large wire, so as to have low resistance, and are connected in series in the circuit (like an ammeter) in which the power is to be measured. The moving coil is made of many turns of fine wire, in order to have high resistance, and is connected in parallel with the load (like a voltmeter). (See Fig. 438.)

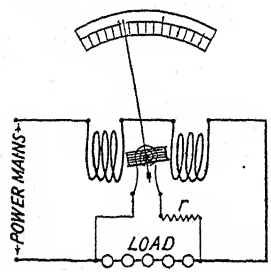


FIG. 438. Diagram of Wattmeter

The flux density B produced by the fixed ("current") coils is then proportional to the current i in the circuit; while the current i'' in the moving coil is proportional to e. The instantaneous deflecting torque $\mathfrak{I}$ is proportional to the product of B and i'', by Ampère's law:

$$\mathfrak{I} \propto Bi'' \propto ie.$$

The **average** deflecting torque $\mathfrak{I}$ throughout a cycle is therefore:

$$\mathfrak{I} \propto \text{average}(ei) = k \text{ average}(\bar{ei}).$$

The restoring torque $\mathfrak{I}'$ due to the spiral springs is, by Hooke's law,

$$\mathfrak{I}' = k'\theta$$

where θ is the angle of deflection.

The moving coil moves into the position in which these two torques are exactly equal, so that

$$k'\theta = k \text{ average}(ei).$$

Therefore,

$$\theta = \frac{k}{k'} \text{ average}(ei). \tag{398}$$

That is, **the deflection of the wattmeter is proportional to the average power throughout the cycle.**

479. Eddy currents. These currents, discovered by Foucault and hence often called "Foucault currents," are produced within the body of a piece of conducting material whenever the magnetic flux through it changes. In accordance with Lenz's law, they are always in such a direction as to oppose the change of flux.

Thus, if a magnet is made to approach a sheet of copper or iron (Fig. 439a), a current will be produced in the sheet as shown. Its direction will be such that it will produce a north pole facing and opposing the approaching N-pole.

If the magnet is moved away from the sheet (Fig. 439b), the eddy current will reverse, so that it produces a south pole which opposes the recession of the N-pole.

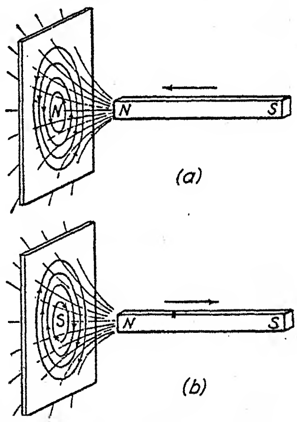


FIG. 439. Eddy, or Foucault, Currents

The phenomenon is shown nicely by suspending a metal disk by a long thread above a horseshoe magnet mounted on a rotator (Fig. 440). When the magnet turns, eddy currents are induced in the disk; and these oppose the relative motion of the magnet and the disk. As the disk is free to rotate, it follows the motion of the magnet below it.

In general, eddy currents are undesirable since their energy is

wasted as heat. They develop in the pole pieces and armatures of generators and motors, and in the cores of transformers. Their effect is minimized by building up such parts out of laminae (thin sheets) of metal, the planes of the laminae being perpendicular to the planes of the currents. In induction motors and watt-hour meters, however, they are made to serve a useful purpose.

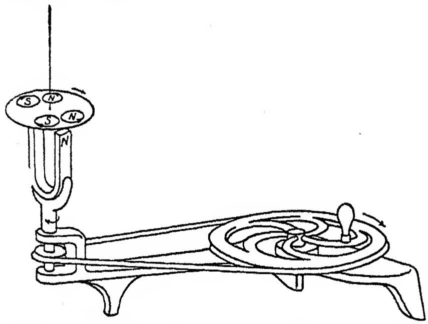


FIG. 440. Demonstration of Eddy Currents

480. The induction motor. Because of the simplicity of its construction, this is the type of alternating current motor in most general use today. Its action depends upon what is known as a **rotary magnetic field,** a phenomenon discovered in 1879 by W. Bailey and developed practically by Professor Ferraris of Turin in 1885.

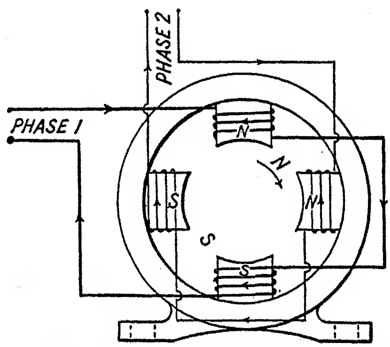


FIG. 441. Stator of Induction Motor

If the poles of a four-pole motor are connected as shown in Fig. 441 so that one pair of opposite poles is energized by the current of phase 1 and the other pair by the current of phase 2, 90° behind phase 1, it will be found that the resultant magnetic field, represented by the large arrow in Fig. 442, will rotate clockwise.

We may trace the poles throughout the cycle and find that the *N*-pole (and of course the *S*-pole) of the resultant field makes a complete revolution during one cycle, just as if a permanent magnet were revolving in the cylindrical space outlined by the actual pole pieces.

Any piece of conducting material placed within this cylindrical space will have eddy currents induced in it in consequence of the rotating field, and will tend to follow the field as did the disk in Fig. 440.

In actual motors, this armature, or "rotor," consists of a cylindrical core of laminated iron in which are embedded stout copper

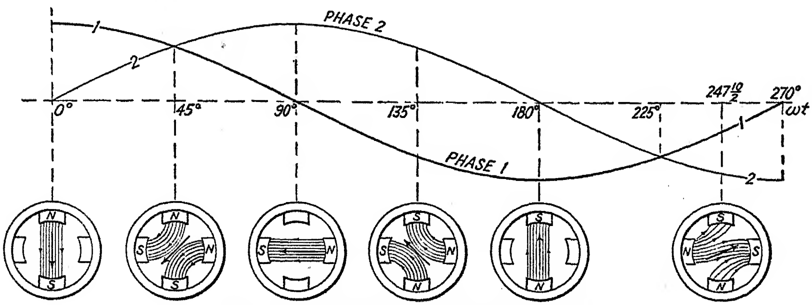


Fig. 442. Rotatory Magnetic Field

rods, insulated from the iron. These are riveted at the ends into copper rings, thus forming the so-called "squirrel cage" rotor (Fig. 443). Large currents are induced in these rods, and since, by Lenz's law, they oppose the change of flux, they tend to follow the rotating field and thus drag the rotor around.

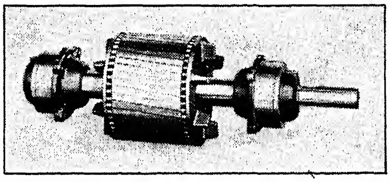


Fig. 443. Rotor of Induction Motor. (Courtesy Allis Chalmers Co.)

If the speed of the rotor were exactly the same as that of the rotating field (synchronous speed), there would be no change of flux through the short circuits of the rotor, hence no induced emf's and therefore no currents. Consequently, the rotor must revolve somewhat less rapidly than the field. This difference between the speed of the revolving field and of the rotor is called the *slip* of the motor, and it increases when the load on the motor is increased.

The direction of rotation of an induction motor may be reversed by reversing the current in either phase.

481. The watt-hour meter. This most extensively used of all electric measuring devices is standard equipment in every building where electrical energy is purchased. It was invented in 1894 by O. B. Shallenberger of the Westinghouse Electric Company. Often erroneously called a wattmeter, it actually measures the product of power and time, i.e., energy, in watt-hours (Fig. 445).

The watt-hour meter is really a small two-phase induction motor with its field poles along a chord instead of arranged around a circle. The frame F (Fig. 444a) is built up of laminations of wrought iron stamped out so as to form the pole pieces of the field magnets. The rotor D consists of an aluminum disk mounted on a vertical shaft J so as to rotate between the pole pieces.

The upper winding has comparatively few turns of large wire, is connected in series in the circuit (like an ammeter), and is called the "current element, or coil." It has small inductance, and the current in it is the line current i or a definite fraction of it. Hence the strength of the upper poles is proportional to i.

The lower coil, or "voltage element," is wound with many turns of fine wire and is connected across the line (like a voltmeter). Its inductance is great; hence its current, proportional to the instantaneous emf e, is 90° out of phase with i (Fig. 444b).

We have here a system of alternate poles excited by a two-phase current and located along a chord of the circular rotor, as shown in the top view (Fig. 444c).

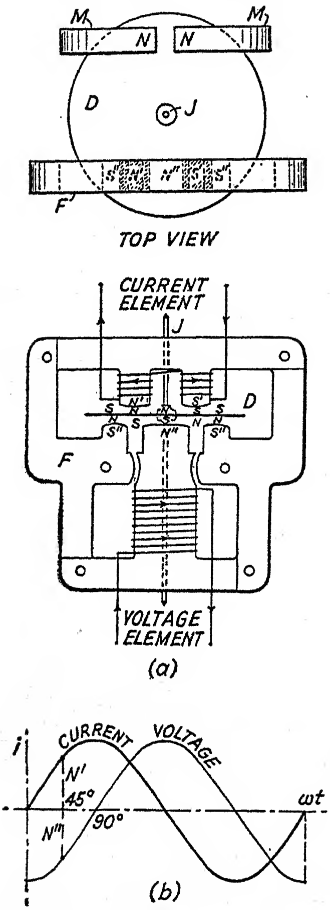


FIG. 444. Mechanism of Watt-Hour Meter

By tracing the polarities of the various poles for a complete cycle of values of both currents, as was done in Fig. 442, remembering that a north pole below the disk is equivalent to a south pole above the disk, it will be found that a north polarity sweeps across the disk from left to right, followed by a south—just as if a permanent horseshoe magnet were moved along a chord from left to right. The eddy currents induced in the disk oppose the motion of these progressive poles, and hence the disk experiences a torque, as shown.

The eddy current induced in the disk by the decreasing S-pole at S'' (left) is proportional to e and is attracted by the increasing N-pole at N', which is proportional to i; and similarly all across the line of poles. Hence the torque $\mathfrak{T}$ on the disk is proportional to the instantaneous power ei, that is

$$\mathfrak{T} \propto ei$$

$$\mathfrak{T} = k(ei). \qquad \text{(a)}$$

At its back edge the disk turns between the poles of two stationary permanent magnets M,M. The eddy currents induced by these magnets react on them and produce a retarding torque proportional to the rate of cutting lines, i.e., to the angular speed.

The disk then comes up to such a speed that the driving torque equals the retarding torque; and this speed is easily shown to be proportional to the power.

The total revolutions (dial readings) are therefore proportional to the product of power by time, or the energy expended in the circuit to which the meter is connected. The dials usually give this energy in **watt-hours** (1 watt-hour = 3600 watt-seconds = 3600 joules).

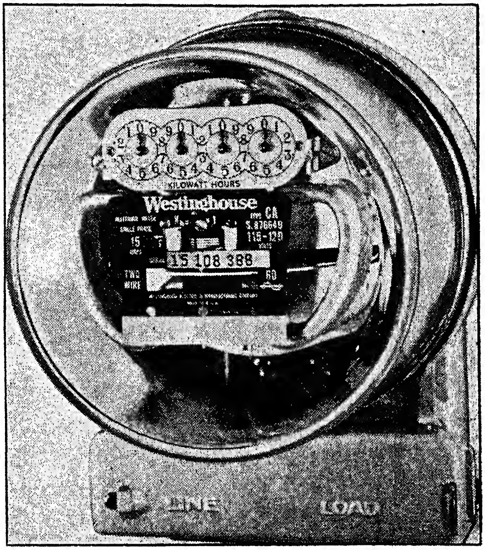


FIG. 445. Watt-Hour Meter. (Courtesy Westinghouse Electric and Manufacturing Co.)

The watt-hour meter (Fig. 445) may be conveniently used to give the average power during a given time by dividing its reading in watt-hours by the time in hours.

482. Synchronous motors. If any two alternating-current generators are brought up to such speeds that they are **in synchronism** (i.e., produce emf's of the same frequency) and are connected in parallel just when they are **in step** (i.e., when they have the same

phase), they will contribute energy jointly to the line. If then the mechanism that drives one of them as a generator is disconnected, that one will cease to act as a generator but will continue to run as a motor with its speed unchanged, receiving energy from the other generator.

Any motor that runs at such a speed (**synchronous speed**) that its frequency as a generator would be the same as that of the impressed emf is called a **synchronous motor.** This is the case if a conductor of its armature just passes one pair of field poles during each cycle of the impressed emf.

Synchronous motors continue to run at synchronous speed as the load on them is increased until they become overloaded, when they stop altogether. The ordinary induction motor may be converted into a synchronous motor by milling in the surface of its rotor, parallel to its axis, as many broad depressions as there are field poles. These depressions being equally spaced around the circumference, the elevations between them serve as poles to the armature. The motor is then brought up (nearly) to synchronous speed as an induction motor. One phase is then cut out by a centrifugal switch, and the motor continues to run as a synchronous motor.

Electric clocks that operate on the ordinary house-lighting circuit are driven by tiny synchronous motors. They are exceedingly simple. The rotor is a laminated cylinder of wrought iron having an even number of teeth—in the simplest case, two, as shown in Fig. 446. The laminated field magnet F is energized by alternating current, usually of 60 cycles. The rotor is brought up to synchronous speed by being spun by hand, a milled head being provided on the back for that purpose.

In Fig. 446a, the lower tooth of the rotor is approaching the N-pole of the field and is magnetized by induction, becoming a S-pole. Similarly, the other tooth becomes an N-pole. At the same time the magnetizing current is decreasing, so that when the rotor becomes horizontal (Fig. 466b), the field poles are both neutral, or of zero pole strength. The inertia of the rotor will carry it by the horizontal position; and on account of the retentivity of the iron, it will not lose its magnetization entirely. Hence, as the current increases in the reverse direction, making the left field pole a S-pole (Fig. 446c), the rotor is repelled in the same direc-

tion as before. If this analysis is carried out for a complete cycle, it will be seen that the rotor makes exactly one revolution during one cycle of the current. That is its synchronous speed.

With 60 cycles per second, there are 3600 cycles per minute. Hence, if the field magnet has two poles, a two-tooth (or pole)

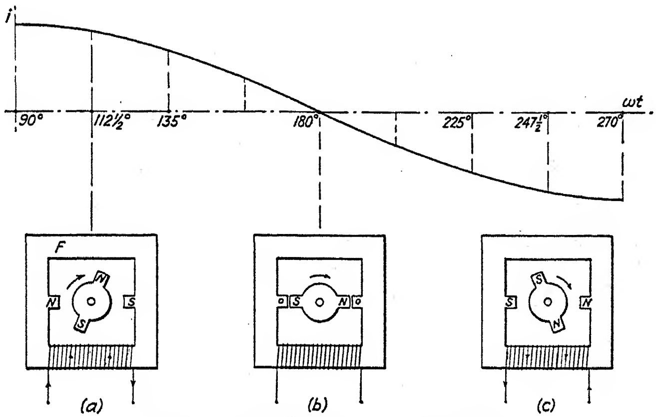


Fig. 446. Synchronous Clock Motor

rotor will have a synchronous speed of 3600 rpm; a four-pole rotor will run at 1800 rpm; an eighteen-tooth rotor, at 400 rpm; and so on. Obviously, hand-started motors will run in either direction.

Self-starting synchronous and induction motors have a second winding of greater inductance than the main winding of the field, connected in parallel with the main winding. This "split-phase" arrangement gives the effect of a two-phase induction motor until the rotor comes up to synchronous speed, after which the second winding is cut out and the motor continues to run as a synchronous motor on one phase. In self-starting clocks the same effect (i.e., a rotating magnetic field) is secured by means of a "shading coil" of heavy copper around the leading tip of each pole of the stationary magnet.

The value of these electric clocks as timekeepers obviously depends upon the accuracy with which the proper frequency of the alternating current is maintained at the power house.

483. The universal motor. Any d-c series motor having both armature and pole pieces laminated will run also on a-c, since both the armature current and the field current are reversed at the

same time. But in general the efficiency will be less on a-c, and there will be excessive sparking at the brushes.

The so-called **universal motor,** for service on either d-c or a-c, is of the above type, but it is more liberally designed with regard to both copper and iron than for d-c only. It should be provided with a compensating winding to reduce the inductance of the armature, and with "resistance leads" to reduce the sparking. Universal motors are very useful for small powers where speed control is important, as for sewing machines and stirrers; they are used for driving railway trains as well.

Direct-current motors with laminated pole pieces will run also when shunt-connected, provided the impedances have proper values, but this is unusual.

484. The transformer. The fact that alternating current is now used for the great majority of industrial purposes is due to the ease with which it may be transformed from high voltage and low current to low voltage and high current, or vice versa. This is accomplished by a *transformer*, the ordinary type of which is shown diagrammatically in Fig. 447 and in cross section in Fig. 448.

It consists of two separate coils having N_1 and N_2 turns, respectively, and wound on a core of iron, laminated to reduce eddy currents. One coil, called the **primary,** is connected to the source of alternating current and **receives energy** therefrom. The other, called the **secondary,** is connected to the load and **delivers energy** to it.

If the secondary voltage is greater than the primary voltage, it is a **step-up** transformer; if the secondary voltage is less than the primary, it is a **step-down** transformer.

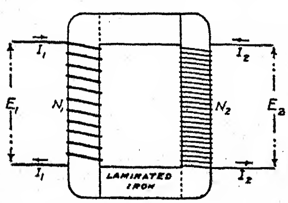


FIG. 447. Diagram of Transformer

The primary current I_1 produces an alternating magnetic flux Φ in the core. This flux interlinks with both the primary and the secondary coils, except for a small **leakage flux** which evades some of the turns. Being approximately sinusoidal, the flux as it varies induces a **back emf** E_1 in the primary which is nearly equal to the emf E impressed on the primary from the power mains, the difference be-

ing just enough to overcome the resistance of the primary and maintain the current I_1 sufficient to produce the flux in the core.

At the same time, this flux, changing through the turns of the secondary, induces in them an emf E_2. When the secondary current I_2 is zero, the primary current I_0, which is called the exciting current, is extremely small. As load is thrown on the secondary, I_2 increases and tends to demagnetize the core. This reduces the back emf in the primary slightly and the primary current I_1 increases until equilibrium is restored. The transformer is therefore a self-governing device.

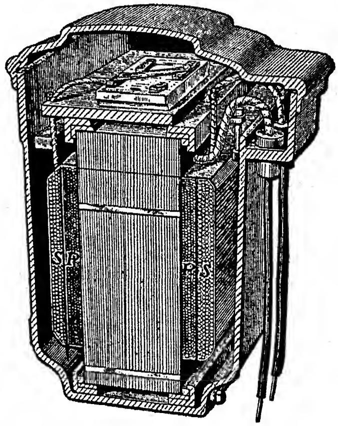


FIG. 448. Cross-Section of Transformer. (From Hoadley's *Essentials of Physics*, American Book Company)

There is no wire connection between the primary and secondary windings; **the energy is transferred from the former to the latter by the alternating flux in the iron core.**

An **ideal transformer** may be defined as a transformer in which there are no losses due to resistance, eddy currents, or hysteresis; no flux leakage; and in which the exciting current is zero. Its efficiency would be 100%. It is not possible, of course, to construct such a transformer, but the concept is useful in theoretical discussion.

Consider such an ideal transformer, and apply the law of Henry and Faraday to its two windings.

In the primary,

$$e_1 = -N_1\frac{d\Phi}{dt}. \qquad (a)$$

In the secondary,

$$e_2 = -N_2\frac{d\Phi}{dt}. \qquad (b)$$

Since in an ideal transformer there is no flux leakage, $d\Phi/dt$ is the same for both coils. So, dividing (a) by (b),

$$\frac{e_1}{e_2} = \frac{N_1}{N_2}. \qquad (c)$$

and since Eq. (c) is true at every instant, it is true for their effective values, so that

$$\frac{E_1}{E_2} = \frac{N_1}{N_2}. \tag{d}$$

In the ideal transformer the efficiency is 100%; hence at every instant

$$\text{Power in} = \text{Power out.}$$

That is,

$$e_1 i_1 = e_2 i_2 \tag{e}$$

which, with Eq. (c), gives:

$$\frac{e_1}{e_2} = \frac{i_2}{i_1} = \frac{N_1}{N_2}. \tag{f}$$

Hence,

$$\frac{I_2}{I_1} = \frac{N_1}{N_2}. \tag{g}$$

Combining Eqs. (d) and (g), we get the **general law of the ideal transformer:**

$$\frac{E_1}{E_2} = \frac{N_1}{N_2} = \frac{I_2}{I_1}. \tag{399}$$

From this,

$$E_1 I_1 = E_2 I_2. \tag{h}$$

But for a-c power in general, we know from Sec. 476 that for 100% efficiency

$$E_1 I_1 \cos \phi_1 = E_2 I_2 \cos \phi_2.$$

Comparing this with Eq. (h), we see that in an ideal transformer,

$$\cos \phi_1 = \cos \phi_2. \tag{i}$$

Hence the primary and secondary currents in an ideal transformer have the same lag behind their emf's.

Alternating-current transformers are among the most efficient devices in general use. Their efficiencies run as high as 93% at one-fourth load to 99% at full load. For this reason, the foregoing relations, though **strictly true only for an ideal transformer,** may generally be applied to actual transformers without serious error.

It should be clearly understood that transformers of these types

cannot be used for direct current, since it is necessary to have the magnetic flux in the core vary. The magnetic flux does not vary with steady direct current.

485. The autotransformer. The autotransformer differs from an ordinary transformer only in having its primary and secondary coils parts of the same winding instead of separate and distinct coils. It is therefore not safe for high voltages, but is very convenient for low transformation ratios.

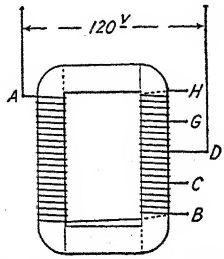


FIG. 449. Diagram of Autotransformer

Equation (399) holds also for autotransformers.

Thus in Fig. 449, let us say there are 800 turns all together (400 on each leg of the core), with taps brought out from points 100 turns apart, as at *B*, *C*, *D*, *G*, and *H*.

Then, if 120-volt mains are connected to *A* and *D* across 600 turns (i.e., 5 turns per volt), we would get voltages as follows:

From *B* to *C*—100 turns— 20 volts
B to *D*—200 turns— 40 volts
B to *G*—300 turns— 60 volts
A to *B*—400 turns— 80 volts
A to *G*—700 turns—140 volts
A to *H*—800 turns—160 volts
etc. etc.

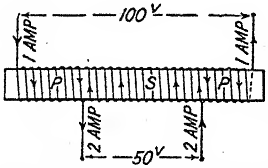


FIG. 450. Simplified Autotransformer

In the simplified diagram of Fig. 450 (the straight core is not efficient) the parts of the winding serving as primary and secondary are marked *P* and *S*, respectively, and the directions of the currents are shown at a given instant.

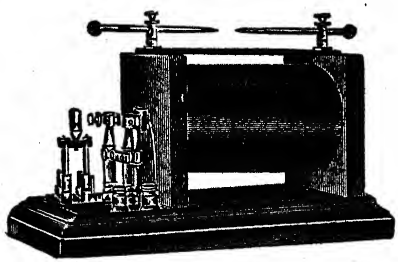


FIG. 451a. Induction Coil. (Courtesy Central Scientific Co.)

486. The induction coil (Fig. 451a) often called the Ruhmkorff coil but actually invented by C. G. Page of Harvard College in 1838, is a transformer for changing low-voltage **direct** cur-

rent into high-voltage alternating current (not sinusoidal). The primary consists of comparatively few turns of large wire wound on a core of soft iron. Over this are wound thousands of turns of fine wire forming the secondary (Fig. 451b).

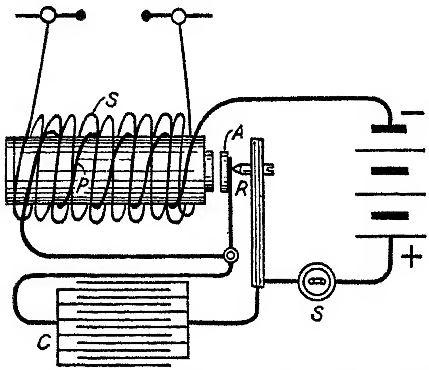


FIG. 451b. Diagram of Induction Coil. (Courtesy of American Book Company)

When the primary is connected to a battery or other source of d-c, the core is magnetized, and the soft-iron armature A is attracted to the left. This **breaks** the contact at R, and the primary current vanishes. There being then nothing to maintain the flux, that disappears also. The armature A is therefore released, springs back, and **makes** contact again at the point R; and the primary current and the flux again build up.

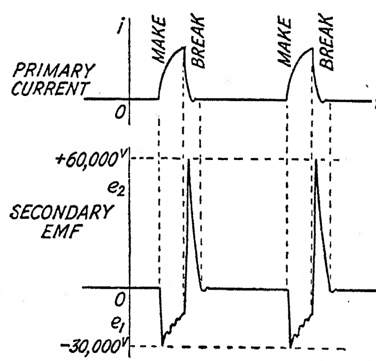


FIG. 452. Curves of Induction Coil

This varying flux through the large number of turns of the secondary induces in that coil a very high emf, which alternates, but not according to a sine law. A condenser C connected across the "gap," and usually concealed in the base of the apparatus, reduces the sparking at the gap and increases the rate of change of flux, and consequently the secondary emf.

As shown in Fig. 452, the primary current dies away on "break" (when the circuit breaker opens) more rapidly than it grows on "make." Consequently the secondary emf is greater on break than on make.

487. The Tesla transformer. From the law of Henry and Faraday,

$$e = -N\frac{d\Phi}{dt},$$

it is clear that large emfs may be secured either by having many turns N or by a great rate of change of flux $d\Phi/dt$. The former method is used in the induction coil; the latter has been employed by Nikola Tesla for the production of exceedingly great emf's of high frequency.

It was shown by Henry and Lord Kelvin that the spark of a condenser is oscillatory with a frequency of the order of magnitude of 1,000,000 cycles per sec. Accordingly, Tesla charges a condenser C (Fig. 453) by means of an induction coil or ordinary transformer T until the air in the spark-gap G breaks down and a spark starts. The condenser then discharges, producing a current of very high frequency through the few turns of the primary coil P. This induces in the secondary S an extremely high emf of the same frequency as the oscillations in the spark discharge. There must be no iron in the Tesla coils P and S.

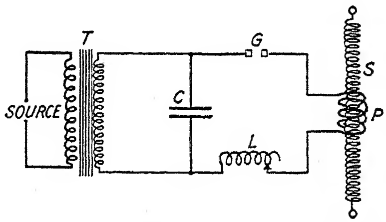


Fig. 453. Diagram of Tesla Transformer

This arrangement is known as a Tesla transformer. By adjusting the inductances of the coils and the capacitance of the condenser until the primary and secondary are in resonance, sparks 8 or 10 ft long and emfs of the order of 1,000,000 volts may be obtained.

At high frequencies electricity does not permeate the entire cross section of a conductor, but flows principally in a thin surface layer. This phenomenon is known as the **skin effect.** Medical treatments are therefore safely given with high-frequency currents, but should not be attempted by untrained persons.

Many striking experiments may be made with high-frequency currents. One may light an electric bulb or a Geissler tube by holding it in the hand and near one terminal of the secondary. The "electric queen" at the circus, who permits the "professor" to draw long sparks from her ears and finger tips and to light a cigarette (held in her lips) by a spark from his wand, is merely sitting on an insulated chair in contact with one terminal of a Tesla transformer, and is in no more danger than is the sea lion on the stepladder.

PROBLEMS

1. A coil of wire has a resistance of 0.8 ohm and an inductance of 50 millihenries. How much energy is stored in its magnetic field when the current is 5 amp?

2. A coil has a resistance of 0.8 ohm and an inductance of 50 millihenries. What current will be produced in this coil by an effective a-c potential difference of 110 volts at 60 cycles?

3. Five kilowatts are received by a motor from a power line each line of which has a resistance of 0.08 ohm. How much power is lost in heating the line when the voltage at the motor terminals is 100 volts? 500 volts?

4. Eleven kilowatts are delivered from a power house to a mill and each wire between the two places has a resistance of 4 ohms. What is the line loss in watts when the power is delivered at 2200 volts? At 22,000 volts? What is the percentage power loss in each case?

5. A generator whose terminal voltage is 220 volts produces a current of 10 amp in the circuit to a motor. If each wire of the power line has a resistance of 0.4 ohm, how much power is wasted in the line and what is the terminal voltage at the motor?

6. A transformer is to step 110 volts down to 6 volts. If there are 550 turns on the primary, how many must be on the secondary? What is the primary current when the secondary current is 5 amp?

7. A transformer is to be constructed to step 110 volts up to 2200 volts. If there are 440 turns on the primary, how many must there be on the secondary? What will be the primary current when that in the secondary is 0.5 amp?

8. A transformer is to step 110 volts down to 20 volts. If there are 550 turns on the primary, how many must be on the secondary? What is the primary current when the secondary current is 10 amp?

9. An a-c voltmeter and ammeter read 110 volts and 6 amp, respectively, when a wattmeter indicates 500 watts. What is the power factor of the circuit and the lag of the current behind the voltage?

CHAPTER XXXV

IMPORTANT ELECTRICAL DEVICES

Besides motors, generators, and the instruments for electrical measurement already mentioned, the increasing number of electrical devices is so great that only a few of the fundamental ones can be described here.

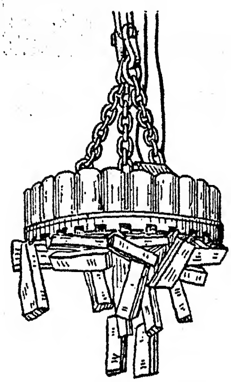


Fig. 454. Lifting Magnet. (Courtesy of American Book Company)

488. Lifting magnets. Electromagnets of great strength are much used in the iron industries for moving unwieldy or rough pieces such as steel beams and pig iron (Fig. 454).

The force F which a magnet exerts upon a piece of ferromagnetic material is given by the relation: *

$$F = \frac{B^2 A}{8\pi} \text{ dynes} \qquad (400)$$

where B lines per cm^2 is the flux density between them and A cm^2 is the area of practical contact.

The magnetic separator. Magnetic material is readily separated from nonmagnetic by the device shown in Fig. 455. The mixed materials are distributed in a thin layer on a belt conveyor B, which passes over a pulley P provided with strong magnets whose poles are in the surface of the pulley. These magnets cause the magnetic material to follow the belt farther than do the nonmagnetic particles, and hence a separation is effected.

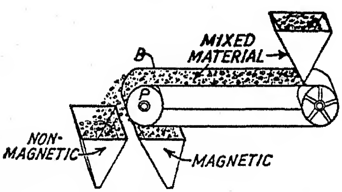


Fig. 455. Magnetic Separator

Surgical magnets are strong electromagnets having a very small pole. By means of them small bits of magnetic material may be

* For derivation see Christie, *Electrical Engineering* (New York, McGraw-Hill Book Co., 1917), p. 68.

withdrawn from the eye or other parts of the body, if not too deeply embedded.

The fact should be carefully noted that magnets exert appreciable attraction only upon iron, cobalt, nickel, and a few alloys, and then only at very short distances. Consequently, gold or silver coins and jewelry cannot be recovered by the use of magnets.

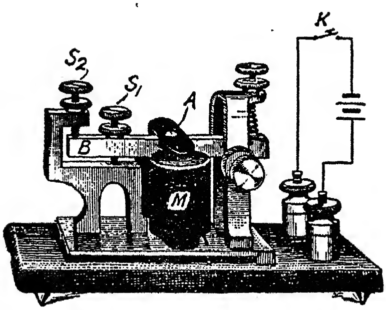


FIG. 456. Telegraph Sounder. (Courtesy of American Book Company)

489. The telegraph. The telegraph (Fig. 456), invented by S. F. B. Morse in 1837, consists of two electromagnets which attract a soft-iron armature A when the circuit is closed by a key K. The armature actuates a bar B, and a click is heard when this bar strikes either of the set screws S_1 or S_2. A short interval between these two clicks is the "dot," and a long interval, the "dash," of the Morse code.

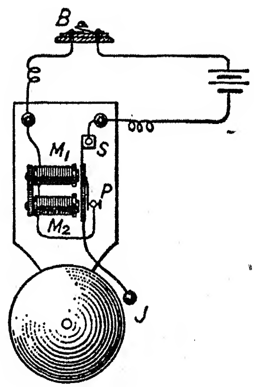


FIG. 457. Electric Bell. (Courtesy of American Book Company)

490. The electric bell and horn. The mechanism of an electric bell is much the same as the Neef magnetic circuit breaker of the induction coil. When the circuit is completed at the push button B (Fig. 457), the electromagnets M_1 and M_2 attract the soft iron armature S, causing the striker J to tap the bell. This motion of S also opens the circuit at P, causing the magnets to lose their magnetism and release S, which is thrown back in contact at P by a spring; and the cycle is then repeated.

In a magnetic horn, instead of having a striker, the armature is attached by a link to the center of a thin metal diaphragm, whose vibration causes the sound.

491. The telephone. A simplified diagram of a telephone circuit is shown in Fig. 458. Sound waves enter the transmitter at the

left and cause diaphragm D_1 to vibrate. D_1 is connected to the movable, left side of the **transmitter button.** This button consists of two plates of solid carbon (shown black), separated by a layer of granulated carbon C.

It is a well-known fact that **the electrical resistance of pieces of carbon in contact decreases when the pressure increases, and vice versa.**

Consequently, as the plate of solid carbon at the left is pushed in and out by the vibrating disk D_1, the resistance of the local

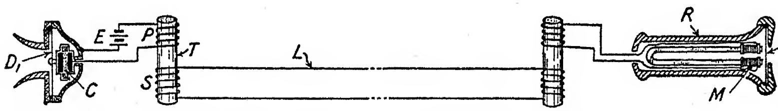


FIG. 458. Telephone Transmitter and Receiver

battery circuit is varied, and the battery E produces through the primary P of the transformer T a current which fluctuates to correspond with the sound waves.

In the transformer the voltage is stepped up and the secondary current is transmitted through the line L. At the receiving end, the voltage is stepped down again and the secondary current passes through the coils M on the two ends of a U-shaped permanent magnet in the receiver R. The varying current in the coils M causes the magnetic pull on the soft-iron diaphragm D_2 to vary. D_2 is thereby made to produce vibrations which are reproductions of the vibrations of the original sound.

Alexander Graham Bell's original telephone (1875) consisted essentially of two receivers similar to R; consequently, it was serviceable only for short distances. The carbon button, devised by Thomas A. Edison in 1877, increased the amount of energy that could be impressed into the line current and thereby increased the useful range of the instrument.

492. The arc light. Sir Humphry Davy was the first to use electricity (1801) for the production of light. He connected two rods of charcoal (carbon) in series with a battery of 2000 cells; and when the carbons were separated from contact, the hot ionized gases between the electrodes formed a luminous arc. Subsequently, various mechanisms were devised for feeding up the

carbons as they burned away. Figure 459 shows a hand-feed carbon arc for use in a projection lantern.

With pure carbon electrodes the arc itself is only slightly luminous, most of the light coming from the "crater" of the + electrode. The luminosity of the arc is greatly increased (flaming arc) by boring out the carbons and putting in a core of metallic oxides and alkali salts.

Arc lights may be operated on either direct or alternating current, but with direct current the arc is less noisy and the luminous

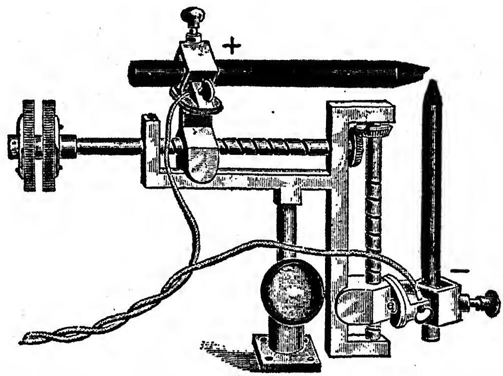


FIG. 459. Arc Light. (Courtesy Central Scientific Co.)

efficiency is higher. This runs as high as 75 lumens, or 6 candles, per watt. However, on account of the deep shadows cast by the light and the difficulty of keeping the lamps in order, they are now little used except for outdoor lighting and searchlights.

The temperature in the crater of the + carbon is one of the highest that can be produced artificially on a large scale, ranging from 3800°K at 1 atmosphere pressure to 5890°K at 22 atmospheres (the temperature of the − electrode is 1000° less). In consequence of this fact, one type of electric furnace is essentially a huge carbon arc lamp in which is placed the material to be heated.

493. The incandescent lamp. As early as 1841, an American, J. W. Starr, endeavored to use a wire heated by an electric current as a practical light source. **When a body is heated until it emits light it is said to be incandescent.** Hence such devices were called *incandescent lamps*.

None was successful until 1879, when Thomas A. Edison in America and J. W. Swan in England produced independently very satisfactory lamps by heating a carbon filament enclosed in an evacuated bulb (Fig. 460). The efficiency of these lamps was low (about 3 watts per candle) and the spectrum of the light was deficient in the blue as compared with daylight.

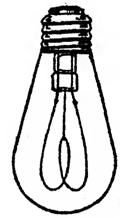


FIG. 460. Carbon Filament Lamp

After years of research, the carbon filament was superseded by one of drawn tungsten wire, developed in 1911 by Dr. W. D. Coolidge of the General Electric Company. Two types of tungsten filament lamps are made.

The **vacuum lamp** (Fig. 461) has its bulb exhausted to a pressure of 0.0001 mm of mercury. Its filament temperature is about 2300°K, and its efficiency is approximately 0.7 candle per watt. It is made only in the smaller sizes. At the high temperature and low pressure, the tungsten evaporates rapidly and coats the inside of the bulb, decreasing the life and efficiency of the lamp.

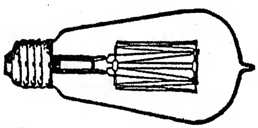


FIG. 461. Tungsten Filament Lamp

To reduce this evaporation, Dr. Irving Langmuir invented the **gas-filled lamp.** In this, the bulb is filled with argon gas mixed with a small proportion of nitrogen, because pure argon will ionize at the temperature used and produce a short circuit between the lead-in wires. The tungsten wire is wound into a small helix with its turns almost in contact. At the higher temperature (2800°K) thus made possible, the spectral quality of the light more nearly approaches daylight, and the efficiency is increased to 1 candle per watt in the 75-watt size and to 1.5 candle per watt in the 1000-watt size.

The candlepower, or intensity, I, of tungsten lamps varies with the voltage E:

$$I = k_1(E)^{3.7} \tag{401}$$

but at the same time the life L of such lamps decreases according to the relation:

$$L = k_2(E)^{-13.8}. \tag{402}$$

494. Discharge tube lights. The so-called "neon" lights are of this type, but only the red ones use neon. Argon in clear glass gives blue; helium in amber glass, yellow; argon and mercury in yellow glass give green. Essentially long Geissler tubes, they are provided with cold electrodes at each end. After complete exhaustion, they are filled with the required gas to a pressure of about 12 mm of Hg and then sealed off. The usual operating voltage is 15,000 a-c, and the power required is small. Satisfactory continuous operation for six years is sometimes secured, but as air leaks in through faulty seals, they become dim and must be re-exhausted and refilled.

495. Fluorescent lamps. These lamps are the most efficient light sources for general purposes. They consist of thin-walled glass tubes from 1 to 1.5 in. in diameter and from 24 to 48 in. long, having a small filament at each end to serve as a hot cathode, and filled with mercury vapor at low pressure. The operating voltage is 110 or 220 volts. At starting, the cathodes are heated to incandescence and emit electrons which ionize the mercury vapor sufficiently to start the low pressure mercury arc. A magnetic or thermal switch then cuts the starting current from the electrodes.

The high efficiency, which runs from 2.5 to 5 candles per watt, is secured by utilizing the fact that most of the energy of the mercury arc is radiated in the ultraviolet. This radiation is absorbed by fluorescent substances with which the inside surface of the tube is coated. These substances re-emit the energy in visible radiation, which is transmitted by glass.

496. The mercury vapor lamp and rectifier. This lamp, brought out by Peter Cooper Hewitt in 1901, usually consists of a glass tube from 2 to 4 ft long and 1 in. in diameter, suspended in an inclined position. The direct-current type has a single iron anode at the upper end, while the alternating-current type (Fig. 462) has two such anodes. In both types the cathode is a pool of mercury at the lower end. The pressure in the tube when the lamp is not in operation is about 0.001 of mercury.

The lamp is started by tilting until the stream of mercury touches the anode, and then releasing; or by a special induction coil that produces by a spark discharge sufficient ions to start the arc.

The spectrum of mercury (Fig. 587) is strong in blue and greenish yellow, but deficient in red as compared to daylight. Hence these lamps are excellent for photography, but are not suitable

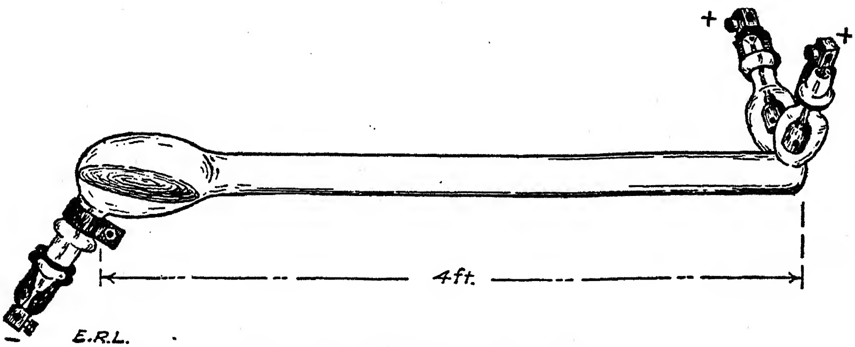


Fig. 462. The Mercury Arc Light

for work in which colors must be matched. Being a large source of low intrinsic brightness (2.3 candles per cm^2), they do not cause heavy shadows. For this reason, and because the spectrum consists of a few strong lines, the strongest being the green (5461 Å), which is in the neighborhood of maximum visibility, mercury vapor lamps are especially adapted for use in industries requiring fine line work, such as engraving and silk weaving. The luminous efficiency of commercial types is of the order of 2 candles per watt.

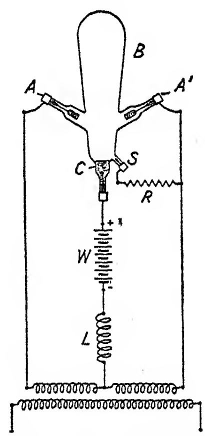


Fig. 463. Mercury Vapor Rectifier

The spectrum of mercury is strong in the ultraviolet, but these short waves are very completely absorbed by glass. They are transmitted, however, by quartz; and if the tubes are made of fused quartz the lamp becomes one of our best sources of ultraviolet radiations.

The mercury vapor lamp was found to have the unusual property of unidirectional conductivity. That is, electrons would pass through it when the mercury electrode was − and the carbon or iron electrode +, but not when the polarity was reversed. In 1903 Dr. Cooper Hewitt successfully employed this property in the construction of the **mercury arc rectifier** (Fig. 463).

Though originally used only for small powers, mercury arc

rectifiers constructed of steel are now used for such purposes as supplying direct current for the New York subway. They are constructed for voltages as high as 16,000, and for currents up to 500 amp.

Their efficiencies range from 80% in the smaller to 98% in the larger sizes, whereas the efficiencies of motor-generator sets run from 46% to 91%.

497. The copper-oxide rectifier. In 1926, L. O. Grondahl * discovered that a disk of copper, having on one side a layer of cuprous oxide (Cu_2O), conducts readily when the direction of the current is from the oxide to the copper, but conducts hardly at all when the direction is from the copper to the oxide. The conductance ratio is of the order of 10,000 to 1 at 20°C, and decreases with rise of temperature.

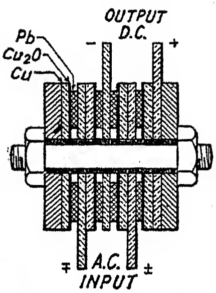


FIG. 464. Copper-Oxide Rectifier

The thickness of the layer of cuprous oxide (red) is not important, as the action seems confined to a microscopically thin layer at the junction of the copper and the crystals of oxide.

Grondahl and Geiger † have developed a dry-plate rectifier employing the above property. It consists of disks of copper about 1.5 in. in diameter, oxidized on one side and assembled upon a bolt, as shown in Fig. 464. The lead washers are used merely to distribute the pressure and to secure good contacts. The bolt is insulated except from the end plates.

The direction of flow of electricity is traced in the diagram of Fig. 465, and it is seen that for both halves of an a-c wave the current on the output side is constant in direction.

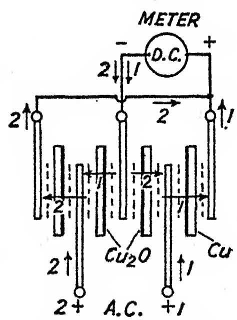


FIG. 465. Currents Traced in Copper-Oxide Rectifier

Since the phenomenon is one of pure resistance, and the leakage is slight, the rectified wave form (Fig. 466) is distorted but little from the input form. This, coupled with the fact that rectification is good up to

* *Physical Review*, Vol. XXVII (June 1926).
† *Journal of the A.I.E.E.* (Feb. 1927).

3,000,000 cycles per second, makes this type of rectifier especially useful in adapting the more sensitive d-c instruments for a-c measurements. Direct-current instruments so used give deflections proportional to **average current** in the instrument; but as this bears a constant ratio to effective value for any given wave form, d-c instruments with built-in rectifiers are calibrated to read effective values. These rectifiers are used largely for charging storage batteries and for radio sets made for a-c operation.

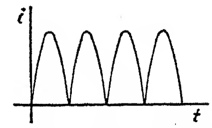


FIG. 466. Full-Wave Rectification of Cu_2O Rectifier

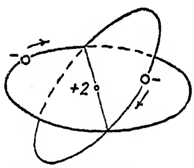
Their efficiencies vary from 65% to 80%; and since neither chemical nor physical change takes place in the constituents, the life of this rectifier is stated in years.

A similar dry-plate rectifier employs disks of cupric sulphide against others of an aluminum-magnesium alloy under pressure of about 200 lb/in.[2] Conduction is from the crystalline sulphide to the alloy, and the action is electrolytic.

CHAPTER XXXVI

ELECTRICAL DISCHARGES IN GASES

498. Ionization. The modified Rutherford-Bohr theory (Sec. 648) assumes that each neutral atom consists of a nucleus made up of a number of protons and neutrons, and that revolving around the nucleus in orbits of various diameters are a number of electrons sufficient to make the complete atom neutral. That is, the nucleus corresponds to the sun and the electrons to the planets, the latter being held in their orbits by the positive charge of the nucleus. The number of these electrons (which is the same as the number of positive charges in the nucleus) equals the **atomic number** of the element, that is, the ordinal number of its position in the periodic table. Figure 467 shows such a hypothetical model of an atom of helium whose atomic number is two. The orbits of the planetary electrons are very large as compared to the size of the proton or the electron, so that it is not difficult to disengage an electron from an atom. (This preliminary picture of an atom is undergoing modifications which will be mentioned in Chap. XLVII.)



BOHR'S HELIUM ATOM
(ATOMIC No.2)

FIG. 467. Bohr's Helium Atom

An **ion** has been defined as an atom or group of atoms that has an excess positive or negative charge. Hence any atom that has **lost** an electron is said to be **positively ionized;** and one that has **gained** an electron is said to be **negatively ionized.**

499. Conduction through gases.* We have seen that, in a metal, conduction consists in the movement of the comparatively "free" electrons through the conductor, and that in a liquid it is due to the migration of ions through the liquid. In fact, to say that a liquid conducts electricity is equivalent to saying that it is partially ionized. If a gas contain ions, we should expect it to con-

* K. T. Compton and I. Langmuir, *Review of Modern Physics*, Vol. II, No. 2 (April 1930).

duct; for the + ions would be attracted to the − terminal and the − ions to the + terminal (Fig. 469), and this flow of electricity would constitute a current. Precisely this is found to take place.

Under ordinary circumstances gases are very poor conductors because they contain few ions. Air, for example, at standard conditions contains only about 20 ions per cm^3.

Warm, damp air is not a better conductor than cool, dry air, as is often supposed. Electrostatic instruments lose their charges

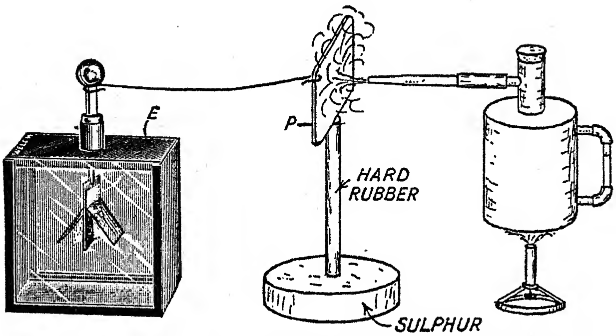


FIG. 468. Moist Air Is Not a Good Conductor. (Courtesy Welch Scientific Co.)

rapidly in warm, damp weather because of leakage due to a film of water on the insulating supports. If these are heated until the film evaporates, the leakage stops.

This can be shown as follows. Let a metal plate P (Fig. 468) be insulated on thoroughly dry sulphur or hard rubber, attached to an electroscope E, and the system charged. The rate of discharge will then be the same, whether cold dry air, hot air, or steam is blown against the plate.

Ionization of a gas may be produced by either of two methods:

I. *By collision, or impact.* That is, electrons are detached from ions and molecules by the impact of other ions or molecules.

II. *By the absorption of radiation* of short wave lengths until an electron becomes expelled from an atom (photoelectric effect).

The practical ways of carrying out these methods in the laboratory are:

Method I—

1. By passing the gas through a flame.
2. By passing the gas through an electric arc.

3. By a spark discharge through the gas.
4. By impact of alpha-particles, beta-particles, deuterons, etc.

Method II—

1. By irradiation with ultraviolet radiation.
2. By irradiation with x-rays.
3. By irradiation with gamma-rays.

The spontaneous ionization that takes place in gases is thought to be due to exceptionally violent impacts of molecular agitation, to small traces of radioactive elements, and to cosmic rays.

500. Ionization chamber, saturation current, and saturation voltage. Quantitative measurements of ionization are made by means of some form of ionization chamber, one of which is shown in Fig. 469. Inside a lead housing H are two metal plates A and B, well insulated from the housing. These are connected in series with a battery of 500 to 800 volts and a quadrant electrometer or a high-sensitivity galvanometer. Ionization currents are of the order of 10^{-13} ampere.

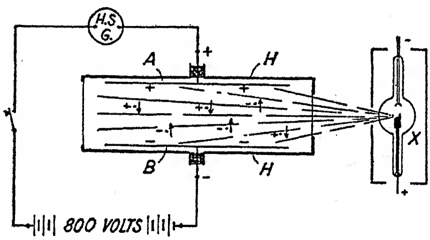


FIG. 469. Ionization by X-Rays

The ionizing radiation, in this case x-rays, is admitted, as shown, through an opening, which may be covered with thin aluminum or Cellophane.

When the x-rays are turned on, the gas between A and B will be partially ionized. The + ions will be attracted to B and the − to A, neutralizing part of the charges on these plates and reducing their potential difference. But the battery at once reestablishes the potential difference between A and B by withdrawing electrons from A and delivering them to B. This stream of electrons passes through the galvanometer and constitutes the **ionization current** which the instrument measures.

At low potential differences between the plates many of the ions that are formed will recombine into neutral molecules before reaching A or B. As the potential difference is increased, a larger and larger number of the ions will be caught by the plates, until finally all that are produced will be caught. The current will then

necessarily become constant for further increase of the potential difference between A and B (Fig. 470).

The ionization current obtained when all the ions produced between the plates are caught by the plates is called the **saturation current**; and the smallest potential difference at which the saturation current is obtained is called the **saturation voltage.**

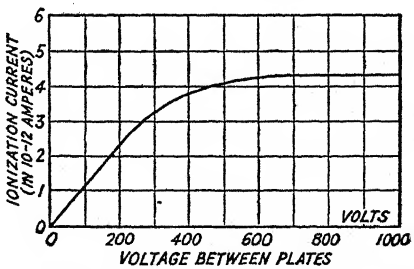


FIG. 470. Saturation Current

501. Discharge in gases at ordinary pressures. Although gases at standard conditions are very poor conductors, we have seen with both the influence machine and the induction coil that if the potential difference between two electrodes is sufficiently high, a bluish glow (**corona**) is visible around both electrodes. This is because of the ionization by impact of the gas near the electrodes. Still further increase of potential difference will cause a crackling, zigzag spark to pass between the electrodes—a miniature lightning flash (Fig. 471).

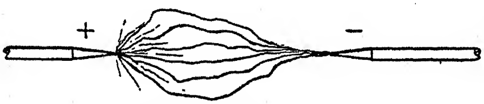


FIG. 471. Spark Discharge between Points

The potential difference at which the discharge takes place is directly proportional to the density of the gas and depends upon the length of the gap, the shape of the electrodes, the kind of gas, and other factors. Thus, in air at standard conditions, to produce a spark 25 cm long requires 20,000 volts between needle points, but 50,000 volts between spheres 6.25 cm in diameter.

Steinmetz and Hayden * have found that 320 volts are required to break down the smallest gap in air. For gaps greater than 0.07 cm in the uniform field between parallel plates in air at standard conditions, they give the relation:

$$E = 2.42 + 30L \tag{403}$$

where E is the breakdown voltage in kilovolts, and
L is the gap length in cm.

* "High-Voltage Insulation," *Journal of the A.I.E.E.* (Jan. 1924).

Thirty thousand volts per cm is generally taken as a rough estimate of the breakdown potential gradient in air.

502. Geissler, Plücker, and Crookes tubes. The study of conduction through gases at reduced pressures began in 1853 when Masson of Paris passed the discharge from an induction coil through a Torricellian vacuum. Heinrich Geissler, a glass blower of Tübingen, who about this time invented the mercury vacuum pump that bears his name, became the most expert manufacturer of the experimental tubes containing various gases at a pressure of about 5 mm of Hg. At this pressure they glow most brilliantly, the colors depending upon the kind of gas used and the composition of the glass of the tube. They were often of the most fantastic shapes (Fig. 472), and at the suggestion of Plücker were called *Geissler tubes*.

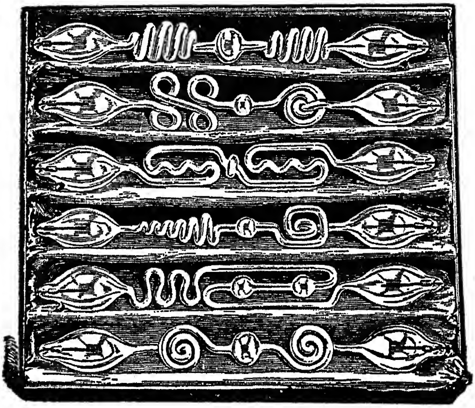


FIG. 472. Geissler Tubes

The many "neon" signs now seen upon our streets are merely enormous Geissler tubes. (See Sec. 494.)

Similar tubes made with a straight capillary at the central section were used by Plücker, Hittorf, and others in the study of spectra. These are commonly called *Plücker tubes* (Fig. 473).

FIG. 473. Plücker Tube

In 1878, Sir William Crookes began the study of conduction in gases. His tubes were of larger sizes and simpler shapes, but exhausted to much lower pressures (of the order of 0.01 mm of Hg) than those of his predecessors. Tubes of this type are generally called *Crookes tubes*.

503. Conduction in gases at reduced pressures. Figure 474 represents a tube 2 or 3 cm in diameter and 30 or 40 cm long, provided with electrodes (usually of aluminum) and connected to a

high vacuum pump. The electrodes are connected to a source of 15,000 to 100,000 volts d-c, or to the secondary terminals of an induction coil.

With air at atmospheric pressure there will be no noticeable discharge. If the pump is now started, it will be observed that when the pressure is reduced to about 100 mm of Hg, faint bluish threads of light appear against the glass. At about 40 mm bright purple streamers connect anode and cathode. As the pressure approaches 10 mm (Geissler vacuum), these streamers unite into

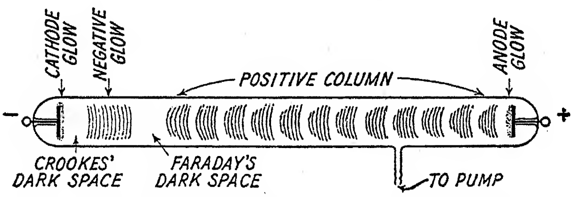


FIG. 474. Electric Discharge in Air at Low Pressure

a steady pink glow, the **positive column.** This extends from the anode nearly to the cathode, from which extends a violet **negative glow,** or column, separated from the positive column by the **Faraday dark space.** At about 4 mm, the brilliant pink column breaks up into striations (buttons), which gradually fade to a white cloud and disappear as the pressure is reduced. At 1 mm pressure (Fig. 474), the negative glow has been separated from the cathode by a second dark band, the **Crookes dark space;** and the cathode glow has appeared on that electrode, this glow being associated with the expulsion of metallic particles from the cathode (cathode sputtering).

As exhaustion proceeds, the Crookes dark space expands until, at about 0.01 mm pressure, it occupies a large part of the tube, which fluoresces in this region with a greenish light due to extremely active ultraviolet radiation from the negative glow. At 0.001 mm (x-ray vacuum), the Crookes dark space fills the entire tube, which now fluoresces with a greenish-yellow light and emits x-rays where electrons repelled from the cathode impinge upon the glass. At 0.000001 mm (Coolidge vacuum) no fluorescence is visible and the tube is nonconducting due to the paucity of ions.

In the tube, the – ions originally in the gas fall toward the

anode, and the + ions fall toward the cathode. As the pressure is reduced, their mean free paths become longer until they acquire enough kinetic energy to produce other ions by collision. In accord with the theory of spectra, the gas then begins to glow on account of the energy radiated when some of the electrons drop back into atoms, or into orbits closer to the nuclei.

The electrons are greatly accelerated in the Crookes dark space, so that on entering the region of the negative glow they are capable of producing ionization by collision (hence the glow). The loss of energy in the collision renders them unable to produce ionization in the Faraday dark space; but they are sufficiently accelerated in that space and in each of the dark spaces between striations to produce ionization (and therefore glow) in the next striation.

The positive ions that do not recombine to form neutral atoms probably do not acquire sufficient energy to produce ionization until they reach the Crookes dark space. They then impinge on the cathode and knock electrons from its atoms, thereby maintaining the supply as the air is removed from the tube.

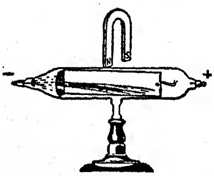


FIG. 475. Cathode Rays Are Deflected by a Magnet. (Courtesy Cambridge Botanical Supply Co.)

When the Crookes dark space has become sufficiently large, the electrons impinge directly upon the glass, causing that to fluoresce with a color characteristic of its composition.

504. Cathode rays. The study of Crookes tubes exhausted to a pressure of about .01 mm engaged many physicists from about 1885 to 1897. At this pressure, the Crookes dark space fills practically the entire tube; but a hazy bluish luminescence seems to mark the path of something which Goldstein called *cathode rays*, emanating from the cathode.*

By means of specially designed tubes, these so-called cathode rays were shown to have the following properties:

1. They ionize a gas, as is shown by the haze marking their path.

2. They are deflected by a magnetic field (Fig. 475).

3. They are deflected by an electrostatic field (Sec. 506).

* J. J. Thomson credits the discovery of cathode rays to Plücker (1859).

4. They have kinetic energy, for they heat an object P on which they impinge (Fig. 476a).

5. They leave a surface normally, because they can be focused by a concave cathode (Fig. 476a).

6. They travel in straight lines, because they cast sharp shadows (Fig. 476b).

7. They produce fluorescence of the glass of the tube (Fig. 476b), and of many other substances.

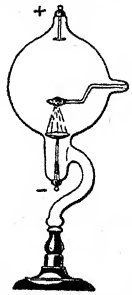


FIG. 476a. Cathode Rays Have Energy. (Courtesy Cambridge Botanical Supply Co.)

These properties are independent of the kind of gas and of the material of the electrodes.

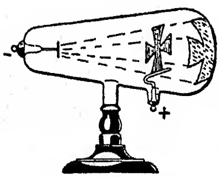


FIG. 476b. Cathode Rays Produce Fluorescence. (Courtesy Cambridge Botanical Supply Co.)

Classic experiments by Sir J. J. Thomson and his students in 1897 proved beyond doubt that **cathode rays are negative electrons.**

505. Thomson's *e/m* experiment. One of the most important experiments in the foundation of modern physics is that of Sir J. J. Thomson for determining the ratio of the charge e to the mass m of an electron, or cathode ray.* His apparatus in principle is shown in Fig. 477.

A thin beam of electrons repelled from the cathode C is selected by slots in the anode A and the screen B and passes through the

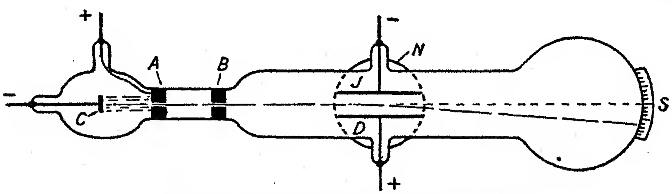


FIG. 477. Thomson's e/m Experiment

strong electrostatic field between the parallel plates D and J. This field alone would deflect the beam of electrons downward, as shown by the long dashes. A uniform magnetic field perpendicular to and out of the paper is produced by external, air-core coils, only one of which, N, is shown. This magnetic field alone would deflect the beam upward, in accord with Fleming's motor rule.

* J. J. Thomson, *Conduction of Electricity Through Gases* (Cambridge University Press, 1906), p. 117.

The electrostatic and magnetic fields are coterminous, so that the length l of the electron path is the same in both. The end of the tube opposite the cathode is coated with a material such as zinc sulphide, which fluoresces at the spot where the beam of electrons strikes it, and a scale S locates the position of this spot.

Let the magnetic field H be turned on first. The beam will be curved upward, and its radius of curvature r may be found from the reading of scale S.

From Rowland's convection experiment (Sec. 408), we know that a moving charge is equivalent to a current of electricity. Suppose there are n electrons in a short length l of the beam, and let an electron traverse this length in time t. Then the quantity of electricity q which passes a point in the path in the time t is:

$$q = ne$$

and from the definition of current I,

$$I = \frac{ne}{t}. \tag{a}$$

The force f on the section of the beam of length l is therefore, by Ampere's law:

$$\begin{aligned} f &= BIl \\ &= HIl \quad \text{in vacuo} \\ &= H\left(\frac{ne}{t}\right)l. \end{aligned}$$

But this f is the force acting on n electrons. Hence the force f_1 acting on each electron is:

$$\begin{aligned} f_1 &= H\frac{e}{t}l = He\frac{l}{t} \\ &= Hev \end{aligned} \tag{b}$$

where v is the speed of the electron.

Since this force is always perpendicular to the path, it is centripetal, and therefore it has also the value:

$$f_1 = \frac{mv^2}{r}. \qquad \text{by Eq. (69)}$$

Hence,

$$\frac{mv^2}{r} = Hev$$

and

$$mv = Her. \tag{c}$$

Let the electric field now be turned on and its intensity F adjusted until the beam of electrons is brought back to its zero position.

From the definition of electric field intensity, the force due to this field on each electron is, by Eq. (257),

$$f_1 = \mathfrak{F}e \tag{d}$$

and since this force neutralizes that due to the magnetic field, we may equate their values from Eqs. (b) and (d):

$$Hev = \mathfrak{F}e$$

$$v = \frac{F}{H}. \tag{e}$$

Substituted in Eq. (c), this gives:

$$m\frac{\mathfrak{F}}{H} = Her$$

whence

$$\frac{e}{m} = \frac{\mathfrak{F}}{H^2r} \tag{404}$$

where all the quantities on the right side are measurable.

The subsequent determination by Millikan, in 1916, of the charge e of one electron made it possible to compute the mass m of an electron. In this way it is found that the

Rest mass of one electron, $m_0 = 9.106 \times 10^{-28}$ gram

which is 1/1838 of the mass of the hydrogen atom. The mass of the electron appears to be entirely electromagnetic, i.e., of the nature of inductance.

506. The cathode-ray oscillograph. Very simple additions to Thomson's e/m apparatus give us the cathode-ray oscillograph tube (Fig. 478a), one of the most useful devices of a modern physics laboratory.

All that is necessary is to add a second set of parallel plates $D2$ and $D2$, and a coating of calcium tungstate or other fluorescent material M on the inner surface of the end of the tube. In the best types, the cathode rays are obtained from a hot, oxide-coated filament used for the cathode, as described in Sec. 509. Enlarged details are shown in Fig. 478b.

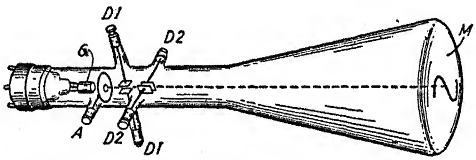


FIG. 478a. Cathode Ray Oscillograph

The fine beam of electrons shot through the hole in the anode is practically a *massless pointer* controlled by the electric fields of the two pairs of parallel plates between which it passes, or by the field of a magnet which may be set up external to the tube.

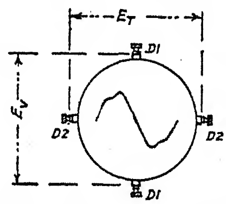


FIG. 478b. End View of Oscillograph

In this way the wave forms of even **radio frequency emf's and currents** may be obtained, since the beam of cathode rays, being without sensible mass, will follow the rapid changes without lag. One of its most recent applications is that of receiving reproducer in television.

507. Positive rays. In the study of cathode rays, one is naturally led to wonder what becomes of the positive parts of the atoms or molecules that have been ionized by the loss of one or more electrons.

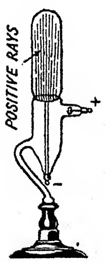


FIG. 479. Positive Rays

As might be expected, these positive ions fall through the potential difference of the tube *toward* the cathode, ionizing the gas as they go and attaining great speed as they pass through the negative glow and the Crookes dark space. They bombard the cathode like a deluge of bullets. It is generally believed that their impacts liberate from the cathode the electrons that maintain the stream of cathode rays, and that **cathode sputtering** is caused by particles of the metal spattered from the cathode by these impacts.

These high-speed positive ions, or *positive rays*, were first detected by Goldstein in 1886, who called them "canal rays." Their

presence is indicated by streamers of light behind the cathode (Fig. 479).

Positive rays are used in the mass spectrograph of Aston, by means of which most of the isotopes have been discovered and Avogadro's law experimentally confirmed.

508. The mass spectrograph; isotopes. In 1910, the English chemist, Frederick Soddy, suggested from radioactive phenomena that elements might exist in two or more forms **having the same chemical properties but differing in mass.** To such forms he gave

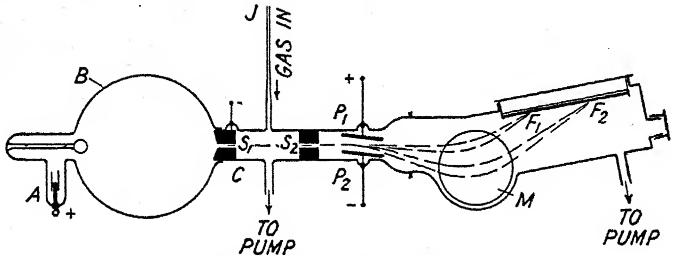


FIG. 480. Mass Spectrograph

the name **isotopes.** Using his parabola method of positive ray analysis,* Sir J. J. Thomson of Cambridge University, in 1913, found evidence of the isotopes of neon.

Certain disadvantages which seemed inherent in the parabola method led F. W. Aston, also of the University of Cambridge, to develop a new type of positive ray analyzer which determined the relative masses of the individual molecules of the gas or gases within the instrument. The apparatus, which he called a **mass spectrograph,** is an outstanding contribution to modern physics.† It is shown somewhat simplified in the diagram of Fig. 480.

The gas to be examined is admitted constantly by the capillary leak J, while the pressure within the apparatus is maintained of the order of 0.005 mm by a vacuum pump. The glass bulb B is about 20 cm in diameter.

A beam of positive ions, repelled from the anode A, is selected by the two slits S_1 and S_2, S_1 being in the cathode C. On passing through the electric field between the charged plates P_1 and P_2,

* *Dictionary of Applied Physics* (London, The Macmillan Co., 1922), II, p. 602.

† F. W. Aston, *Mass Spectrograph and Isotopes* (London, Edw. Arnold and Co., 1933), p. 38.

the ions are deflected through angles proportional to e/mv^2. Passing onward through the field of the magnet M, only one pole of which is shown, they are given a reverse deflection proportional to e/mv, so that all molecules of a given mass are "focused" at

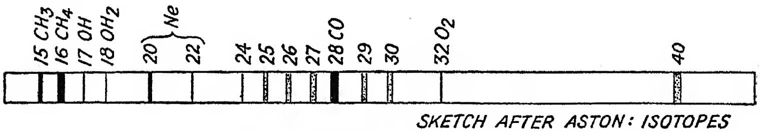


FIG. 481. Mass Spectrogram

some point F_1 on a photographic plate, while those of a greater mass are focused at F_2, etc. The resulting spectrum (Fig. 481) consists of a number of lines, each corresponding to molecules of a certain mass.

In this way the existence of the isotopes (20 and 22) of neon was definitely demonstrated; and most of the isotopes of the other elements have been discovered in the same manner. For his work in this field Aston was awarded the Nobel prize.

CHAPTER XXXVII

ELECTRONICS

509. The Edison effect. In the year 1883, while he was developing the carbon incandescent lamp, Thomas A. Edison placed a plate between the two branches of the filament and observed that a galvanometer which was connected to the plate showed a small current when its free end was connected to the positive terminal of the lamp, but did not do so when it was connected to the negative terminal.

These facts were not satisfactorily explained until 1901, when O. W. Richardson showed that **all hot bodies emit electrons.*** This phenomenon is known as the *Edison effect*, because it is essentially what Edison had observed. It is well shown by the apparatus of Fig. 482.

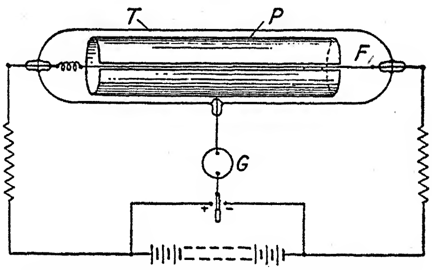


FIG. 482. Tube Showing Edison Effect

The highly exhausted glass tube T contains a cylindrical metal plate P, along the axis of which is stretched a filament F. When the filament is heated to incandescence it emits electrons which are attracted to the plate if that is +, and which constitute a current as shown by the deflection of the galvanometer G. If the plate is connected to the − terminal, the electrons are repelled and G shows no current.

From thermodynamic reasoning, Dushman has shown that the saturation current I_s from a hot filament is given by the relation:†

$$I_s = AT^2\epsilon^{-\frac{b_0}{T}}$$

where A is a constant (about 60.2 amp cm^{-2} deg^{-2} for tungsten);

* O. W. Richardson, *Emission of Electricity from Hot Bodies* (New York, Longmans, Green and Co., 1921), p. 110.

† S. Dushman, "Electron Emission," *Electrical Engineering* (July 1934).

T is the absolute temperature ($°K$);
ϵ is the base of Naperian logarithms;
b_0 is a constant for a given metal (52,400 for tungsten); and
I_s is saturation current in amperes per cm^2.

Langmuir and others have shown that electrons are evaporated from a hot body very much as water molecules are evaporated from the surface of that liquid.

On leaving the hot body, each electron takes from it a definite amount of energy known as the **work function** w for that temperature. It is analogous to the heat of vaporization. The constant b_0, above, depends upon the work function at absolute zero, which for tungsten is about 7.20×10^{-12} erg per electron.

Wehnelt discovered in 1904 that if a platinum filament is coated with an oxide of barium or strontium, its emission of electrons when hot is enormously increased. The best emitters are still these oxide-coated filaments.

The Edison effect is of great importance, because upon it depends the action of most of the vacuum tubes now used in radio, x-rays, etc.

510. Thermions. While positive ions are emitted from hot bodies much less freely and generally than are electrons, it is well known that the alkali and alkaline earth metals, if coated on a filament of a more electronegative metal, will emit ions of their own atoms when the filament is heated; but that they do not evaporate positive ions when heated alone.

Wahlin has found * that chromium, molybdenum, tungsten, and a few other metals, when sufficiently heated, yield positive ions characteristic of these metals as shown by the mass spectrograph; but it appears that iron, nickel, copper, platinum, and others do not.

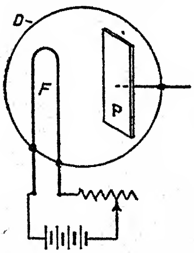


FIG. 483. Diode, or Fleming Valve

The name *thermions* is applied to any ions, positive or negative, that are emitted by heated metals.

511. The Fleming valve. One of the earliest applications of the Edison effect was made in 1896 by Professor J. A. Fleming, in his search for a detector of radio signals. This

* Compton and Langmuir, in *Review of Modern Physics* (April 1930), p. 140.

two-electrode vacuum tube, or **diode**, consists of a hot filament F and a cool plate P, enclosed in an exhausted bulb D (Fig. 483).

Whenever the plate is +, electrons emitted by the filament are attracted to the plate and constitute a current through the tube. But when P is negative, the electrons are repelled from it and there is no current through the tube. The device acts, therefore, like a "check-valve" for electric current, permitting a flow in one direction only. At any instant there is a swarm, or cloud, of electrons called the **space charge** between F and P, and these repel the electrons nearer the filament back toward it, thereby reducing the flow. The space charge plays an important role in all thermionic devices.

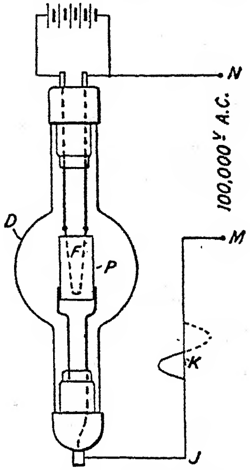


FIG. 484. A Kenotron

As a detector of radio-frequency currents, the Fleming valve was at first only fairly successful, but with improved amplifiers it is coming back into use.

512. The kenotron. This type of diode, developed by Dr. Saul Dushman of the General Electric Company, is the standard rectifier for small currents at high potentials. As shown in Fig. 484, it consists of a hot tungsten filament F surrounded by a cool, cylindrical plate P of tungsten or molybdenum, both being enclosed in a glass bulb D exhausted to a pressure of the order of 10^{-6} cm of Hg.

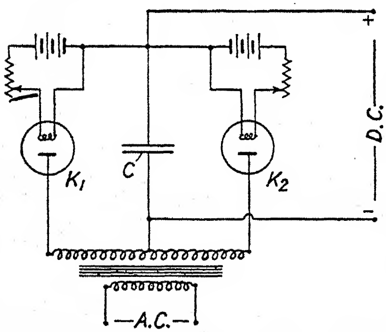


FIG. 485. Connections of Kenotrons for Full-Wave Rectification

At this pressure the gas is so completely removed that it plays no part in the action of the tube, the conductivity of which is due entirely to the electrons emitted by the hot filament (cathode). As long as the plate (anode) remains relatively cool as compared to the filament, electrons are permitted to pass only in the direction from filament to plate.

A single tube is, therefore, a half-wave rectifier, passing only that half of the a-c wave for which the plate is positive.

Commercial tubes are made to rectify currents of the order of 0.3 amp at potentials as high as 140,000 volts.

For full-wave rectification, two kenotrons may be connected, as shown in Fig. 485. For the sake of simplicity, the filaments are shown heated by batteries in these figures. In practice this is usually done by means of small filament transformers.

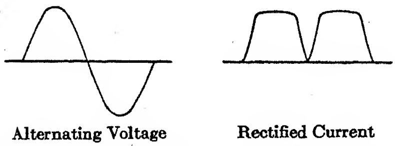


FIG. 486. Full-Wave Rectification by Two Kenotrons

The rectified full wave when the circuit of Fig. 485 is used is shown in Fig. 486. Its trapezoidal form is due to the condenser C.

513. The tungar rectifier.* For rectifying currents of the order of 6 amp at potentials of 100 volts or less, the tungar is used. It differs from the kenotron in that it has a carbon plate P and is filled with the inert gas, argon, at a pressure of about 5 cm of mercury.

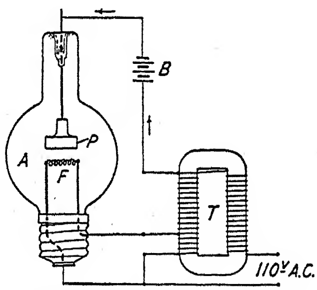


FIG. 487. Tungar Rectifier

Primarily a half-wave rectifier, the tube and connections are shown in Fig. 487. The filament is energized by a section of the autotransformer T. For full-wave rectification, connections similar to those for kenotrons may be used.

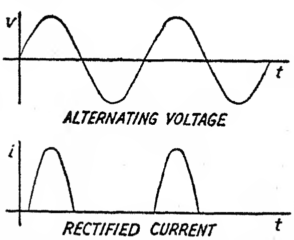


FIG. 487a. Half-Wave Rectification by One Tungar

The rectified wave is shown in Fig. 487a. The efficiency of a tungar rectifier may be as high as 45%.

513a. Full-wave rectifier. The full-wave rectifier (Fig. 488) has a single filament F and two plates P_1 and P_2. The tubes are exhausted to a high vacuum for use at high voltage; but for high current at low voltage, they are filled with mercury vapor.

The input a-c is supplied to an iron-core transformer having

* G. S. Meikle, *General Electric Review* (April 1916).

one primary and two secondary windings. The upper secondary serves to heat the filament, the lower to produce an alternating potential difference between P_1 and P_2.

When P_1 is + with respect to P_2, electrons from F are attracted to P_1. This is equivalent to a flow of positive electricity as shown by the No. 1 arrows. When the potentials are reversed, electrons are attracted to P_2, and the conventional current is shown by the No. 2 arrows. In both cases the polarity of the external d-c circuit is the same and the rectified current has the wave form of Fig. 466.

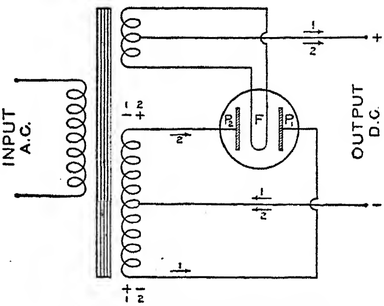


FIG. 488. Full-Wave Rectifier

514. The Coolidge x-ray tube. The earlier types of x-ray tubes were unsatisfactory because when used the vacuum became higher, and changing the potential difference across the tube changed both the quality (wave length) and the quantity (intensity) of the rays in an unpredictable way.

These difficulties were overcome in the type of tube developed by Dr. W. D. Coolidge * of the General Electric Laboratories (Fig. 489). The hot cathode F is a flat spiral of tungsten. The electrons evaporated from this filament are focused upon a small spot on the cold anode P, which is usually a heavy block of tungsten or molybdenum.

The bulb is exhausted to a pressure of the order of 10^{-6} mm of Hg ("Coolidge vacuum").

The electrons acquire enormous speeds as they fall through the high potential difference between F and P. When they impinge upon the anode, or target, some of their kinetic energy is transformed into the energy of very short electromagnetic waves, which we call **x-rays**, and is radiated from the focal spot on the anode.† The rest of the energy of the electrons goes into heat. In the self-rectifying type of tube this heat is dissipated by the radiator fins R. In tubes used for crystallographic work, which

* W. D. Coolidge, *Physical Review* (Dec. 1913).
† This is the inverse photoelectric effect.

requires exposures several hours long, the anode is cooled by circulating water through it.

The quality, or wave length, of the x-rays depends upon the potential difference between F and P: the higher the p.d. the shorter the wave lengths, i.e., the "harder" and more penetrating are the x-rays.

The intensity (the quantity per unit area) of the x-rays depends upon the number of electrons that strike the target per

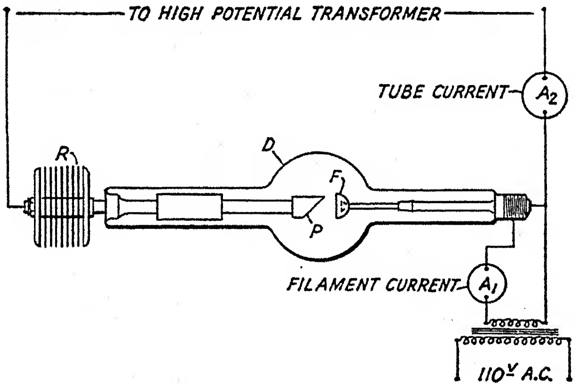


FIG. 489. Coolidge X-Ray Tube

second; and this depends upon the temperature of the filament and hence upon the filament current A_1.

These two factors can therefore be controlled independently, and the performance of the tube is highly satisfactory. It has displaced the earlier gas tube for most purposes.

515. The three-electrode vacuum tube, or triode. The three-electrode vacuum tube was originally called an "audion" by its inventor, Dr. Lee De Forest,* but the name **triode,** suggested by Eccles, is now generally used. It differs from Fleming's valve by having a third electrode in the form of a wire **grid** G between the filament F and the plate P (Fig. 490). The functions of these electrodes are as follows:

The **filament** is heated by the A-battery and **emits electrons** by the Edison effect. (The negative end of the filament is usually taken as the point with reference to which potentials in the system are measured.)

* De Forest: U. S. Patent No. 841,387 (Jan. 1907).

The **plate** is always maintained at a positive potential with reference to the filament by means of the B-battery, and it therefore **attracts electrons,** which form the plate current. (The conventional direction of the plate current is from plate to filament; i.e., opposite to the electron current.)

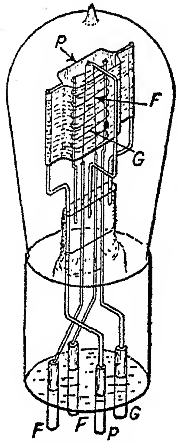


FIG. 490. Triode

The **grid controls the flow of electrons** from filament to plate. When the grid is + with respect to the filament, it attracts electrons and hence increases the plate current; when −, it repels electrons and decreases the plate current. Being closer to the filament than is the plate, its effect is relatively greater than that of the plate. Thus, a very small amount of energy applied to the grid circuit will control a large amount of energy in the plate circuit.

The grid is analogous to the valve in a steam engine: a small amount of energy employed in moving the steam valve admits to the cylinder, or cuts off from it, a large amount of energy.

The **triode has three primary uses,** which are discussed in Secs. 517–519:

1. As a detector of radio-frequency currents.
2. As an amplifier for magnifying currents.
3. As an oscillator, or generator of alternating currents.

516. Characteristic curves. For every thermionic tube there are certain *characteristic curves* which describe the performance of the tube under different conditions. For a triode the most important of these curves is the grid-potential-plate-current curve (Fig. 491). (Potentials are measured with reference to the negative end of the filament.)

On the curve for $E_p = 90$ volts, it will be seen that when the grid-potential E_g is zero the plate-current I_p is 6 ma; and this current is due to the plate-potential E_p alone, since the grid is now neither attracting nor repelling electrons.

As the grid potential is increased the plate current increases until, when $E_g = 11$ volts, **saturation** is reached; i.e., further increase of grid potential produces no further increase in plate

current, because the plate (and grid) are then catching all the electrons that evaporate from the filament at its then temperature.

On the other hand, if the grid potential is made negative, it repels electrons; and for a value of $E_g = -8$ volts, the plate cur-

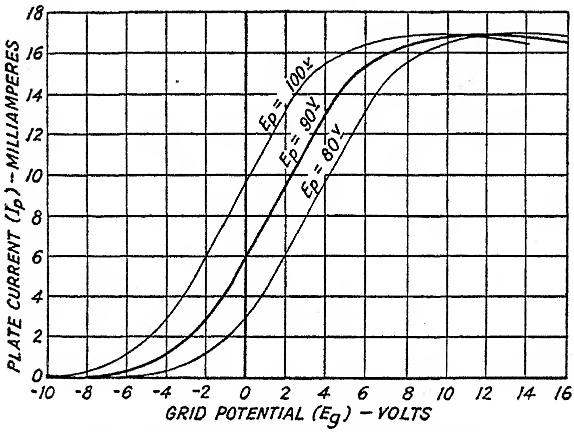


FIG. 491. Characteristic Curves of Triode

rent becomes zero; i.e., the negative grid completely counteracts the effect of the positive plate.

517. The triode as a detector. For use as a detector, which is the purpose for which the triode was originally intended, the simplest connection is that shown in Fig. 492.

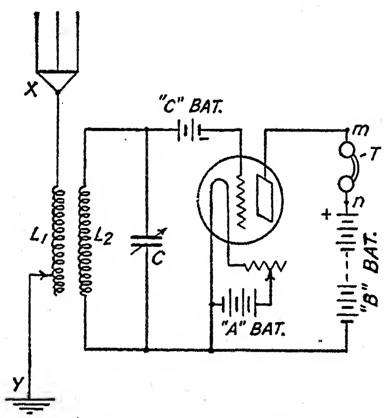


FIG. 492. Triode as Detector

If an alternating emf of radio-frequency (i.e., above 10,000 cycles/sec) is impressed upon the telephones T, practically no sound will be heard on account of the inductance of the coils of the phone and the inertia of its diaphragm.

But if an emf of that frequency is impressed on the grid circuit, the tube will produce a rectifying effect, so that the telephones will respond to each group of waves for which there is a change of amplitude. This will be understood from Fig. 493.

The curve MN is part of the grid-potential-plate-current curve for the tube in question. The curve $abcd$ is the curve of impressed grid voltage. Its time axis Ot is vertical, and its E_g axis OX is horizontal and the same as that for MN. The curve $a'b'c'd'$ is the curve of partially "rectified" plate current. Its time axis Mt' is horizontal, and its plate-current axis OY is vertical and the same as for MN. The same time intervals are laid off along the axes Ot and Mt'.

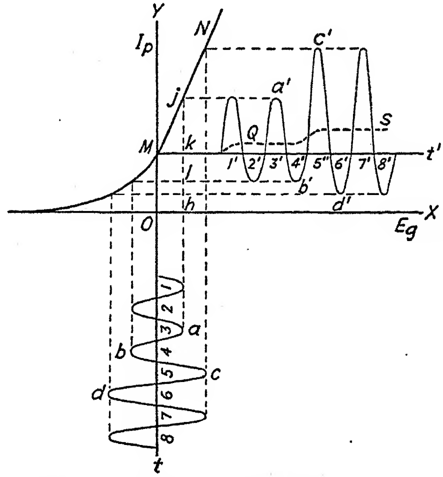


FIG. 493. Curves Showing Detector Action

Suppose the impressed grid potential at a certain instant is represented by $3a$. If we project the point a vertically up to the curve MN, we get hj, which is the plate current at that instant. By projecting j across horizontally to meet the ordinate from $3'$ on the time axis Mt', we find the point on the plate-current curve corresponding to the point a on the grid-voltage curve. All points on the plate-current curve are found in a similar way.

It will be seen that on account of the curvature of MN the positive loops of the plate-current curve are larger than the negative loops, so that the average plate current as shown by the dotted line QS is unidirectional. That is, a certain amount of rectification has taken place.

The telephone disk will respond to this average current for each group of waves whose amplitude is different from the preceding group. Hence the frequency of vibration of the disk is the **group, or speech, frequency,** and not the frequency of individual radio waves.

518. The triode as an amplifier. Since the grid of a triode is closer to the filament than is the plate, a given charge on the grid will be more effective in controlling the plate current than will an equal charge on the plate. Hence the tube

will act as an amplifier, connections for which are shown in Fig. 494.

An alternating voltage is impressed at the terminals m_1n_1 of the input transformer T_1. In the case of radio signals, this may be done by connecting the primary of T_1 in place of the telephones in Fig. 492. (In this case the signals have been made of audio-frequency by the detector, and the transformer may have an iron core; but when radio-frequency currents are amplified, the transformer is usually without iron.)

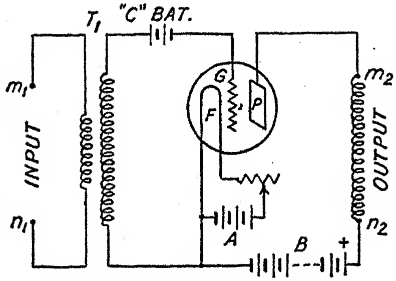


FIG. 494. Triode as Amplifier

The input voltage is stepped up and impressed on the grid circuit of the amplifying tube. To prevent distortion of the wave form, a C-battery of from 3 to 50 volts is inserted in the grid circuit so that the grid will at no time become positive to the filament.

The action of the tube is readily understood from Fig. 495. NW is part of the characteristic curve of the tube. The voltage of the C-battery is adjusted so that when no signal is coming in, the grid will be at the negative potential of the mid-point of a **straight section** NS of the curve, so that none of the signal voltages will fall outside this section.

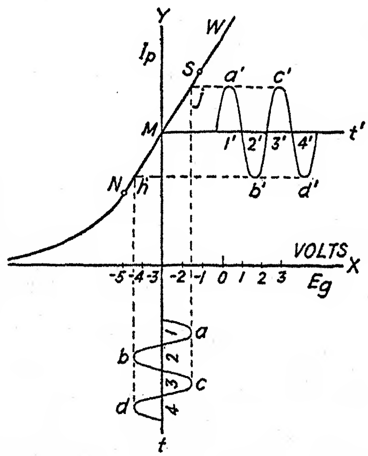


FIG. 495. Curves Showing Amplifier Action

From the input voltage wave $abcd$, the output plate-current wave $a'b'c'd'$ is constructed exactly as in the preceding section; and the output voltage wave will be a magnified facsimile of the input wave, since curve NS is straight.

The amplification factor μ is defined as the ratio of the change of plate voltage ΔE_p necessary to produce a given change of plate current, to the change of grid voltage ΔE_g necessary to produce the

same change of plate current; i.e.,

$$\mu \equiv \frac{\Delta E_p}{\Delta E_g}. \tag{406}$$

In ordinary commercial amplifier tubes such as the 01A type, a 1-volt change of grid voltage produces about the same change of plate current as an 8-volt change of plate voltage produces. Hence the amplification factor for these tubes is about 8. The larger the number of turns forming the grid and the nearer the grid is to the filament, the greater the amplification factor. Tubes have been made with $\mu = 200$.

Amplifying tubes should be "hard," i.e., they should be exhausted to a pressure of 10^{-5} to 10^{-6} mm of Hg.

519. The triode as an oscillator. Alternating currents of frequencies up to 50 kc/sec can be generated by rotating electromagnetic generators, but above that frequency and up to 300,000 kc/sec, the only successful generator is the triode. In its role of generator it is usually called an *oscillator* because it generates only alternating (oscillating) currents. One of the simplest connections for the triode as an oscillator is the **Hartley circuit** (Fig. 496).

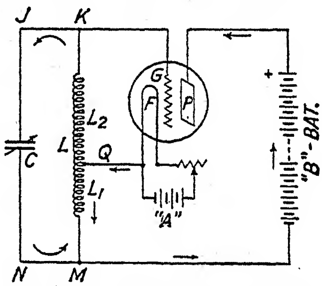


FIG. 496. Triode as Oscillator

An explanation of the action of the oscillator requires setting up the equations connecting the magnitudes of the various currents and voltages, which is beyond the scope of this text. However, the following analogy between a triode oscillator and an ordinary pendulum clock is illuminating.

The clock has a timing device, the pendulum, which determines the frequency of its oscillations. Its period is:

$$T = 2\pi\sqrt{\frac{I}{Mgh}}. \qquad \text{by Eq. (148)}$$

The oscillator has a timing device, the "L-C circuit" $JKMN$, which contains the variable capacitance C and the inductance L. Its period of electric oscillation is:

$$T = 2\pi\sqrt{LC}. \tag{407}$$

The source of energy in the clock is the mainspring; in the oscillator, it is the B-battery. The device in the clock for dealing out energy is the escapement; in the oscillator, the grid.

If a pendulum is set oscillating, it will continue to do so with diminishing amplitude until its available energy is all wasted as heat in overcoming the friction of air and bearings. Similarly, if an electric oscillation (alternating current) is set up in the L-C circuit *JKMN*, the oscillations will continue with diminishing amplitude until all the available energy is wasted as heat in overcoming the resistance of the circuit.

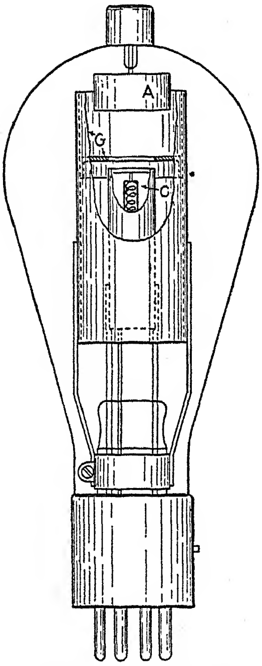


Fig. 497. A Thyratron

In a clock, we maintain the oscillations indefinitely by passing enough energy from the mainspring to the pendulum during each complete vibration to make up for the loss in overcoming friction during that vibration. This is done by means of the escapement. In a similar way, the electric oscillations in the L-C circuit are maintained: during each oscillation (cycle) enough energy is passed from the B-battery to the L-C circuit to make up for the energy lost in overcoming resistance in the circuit during that cycle. This is done by means of the grid.

Thus, suppose the current in the L-C circuit is counterclock-wise, as shown in Fig. 496. The grid is then becoming more positive with respect to the filament; hence the plate current I_p is increasing. This produces a back-emf in the sense (MQ) which aids the oscillation. It also increases the energy stored in the magnetic field of the coil (L_1). When the current in the L-C circuit reverses, the grid becomes less positive and I_p decreases. This develops a back emf in the sense QM, which again aids the oscillation, provided the phase relations of emf's and currents are properly adjusted.

Triodes for transatlantic telephony are made in powers as high as 20 kw. The plates of these high-power tubes are made to form part of the enclosing tube so that they may be water cooled.

520. The Thyratron. This versatile tube is a triode filled usually with mercury vapor at a pressure of from 1 to 100 microns, although argon may be used. Figure 497 shows one form of thyratron. Electrons emitted from a cylindrical hot cathode C are received by the anode A. The purpose of the grid G is to start the arc.

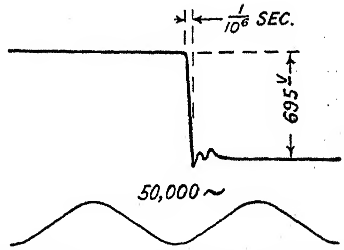


FIG. 498a. Curve Showing Abrupt Starting of Arc

As the potential of G is raised, electrons emitted by C are accelerated and ionize the gas between the cathode and the anode by impact. At a certain critical grid potential, the ionization becomes sufficient to enable an arc to spring from C to A. **The arc starts abruptly** as this critical potential is reached (Fig. 498a). Once started, it cannot be controlled or stopped by the grid; but as it goes out automatically twice per cycle when the plate potential goes through zero, this is no serious disadvantage.

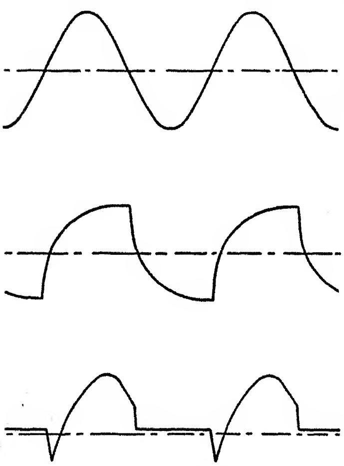


FIG. 498b. Wave-Forms Obtained with Thyratron

The thyratron is used for rectification of a-c; for "inversion" (changing d-c to a-c); for switching and control of machinery by small energy impulses, as from a microphone or photoelectric tube; and for various trick purposes requiring trigger action.

Figure 498b shows three of the many wave forms that may be obtained with various load conditions.

521. The photoelectric effect. While endeavoring to demonstrate experimentally the existence of electromagnetic waves that had been predicted by Maxwell, Heinrich Hertz discovered in 1887 that if the negative terminal of a spark gap were illuminated with ultraviolet light, a longer spark could be induced than without such illumination. Pursuing the matter further, Hallwachs

found (1888) that a polished zinc plate, insulated and connected to an electrometer, exhibited a positive charge after being illuminated with ultraviolet light.

In 1899, Philipp Lenard and J. J. Thomson showed independently that both these effects are due to **the loss of electrons by the body under the action of the light.** The phenomenon is therefore called the surface *photoelectric effect* * (also the "Hallwachs effect"). It is described as follows:

Many substances emit electrons when illuminated with light of the proper frequencies. The effect is most marked with the alkali metals (K, Na, Rb, Cs), their oxides and hydrides.

This fact could not be rationally accounted for by the wave theory of light. However, it fits nicely into the quantum theory (Sec. 616), in accordance with the *photoelectric equation* proposed by Einstein in 1905: †

$$\tfrac{1}{2}mv^2 = h\nu - h\nu_0 \tag{408}$$

where m is the mass of an electron;
v is the maximum velocity of emission;
h is Planck's constant;
ν is the frequency of the incident light; and
ν_0 is the threshold frequency for the substance ($\nu_0 = 5.15 \times 10^{14}$ for Na).

The equation states that when a quantum $h\nu$ of energy is absorbed by an electron, the amount $h\nu_0$ is required to enable the electron to break through the surface of the substance, and that the remainder ($1/2\ mv^2$) is the kinetic energy with which the electron escapes.

Einstein's photoelectric equation has received complete experimental confirmation by Millikan and others. So important has been the role of this equation in the development of quantum theory, that for its conception Einstein was awarded the Nobel prize in 1921.

522. Photoelectric tubes. Many different types of tubes, or cells, have been devised for the study and utilization of the surface

* Hughes and DuBridge, *Photoelectric Phenomena* (New York, McGraw-Hill Book Co.).

† Millikan, *The Electron* (Chicago University, 1935), p. 238.

photoelectric effect. In many of these, the substance that yields the electrons is deposited directly on the inner surface of the glass bulb. One of the more recent commercial forms is shown in Fig. 499. The light-sensitive material is deposited on the concave surface of the metal half-cylinder, which is the cathode; the anode is the metal rod placed along the axis of the cylinder. The bulb is usually of glass, but it must be of quartz if the cell is to be used with ultraviolet light.

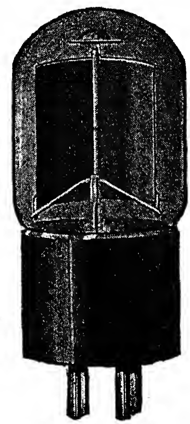


Fig. 499. Photoelectric Tube. (Courtesy of G-M Laboratories)

Photoelectric cells are made in two types:

1. The vacuum (Type AV) cell is exhausted to a high vacuum (10^{-6} mm of mercury). In this type the current is due to the photoelectrons only.

2. The gas-filled (Type A) cell is first exhausted of air and then filled with one of the inert gases —helium, argon, or neon—at low pressure. Within their range, the gas-filled cells yield larger currents than do the vacuum type under the same conditions. This is because

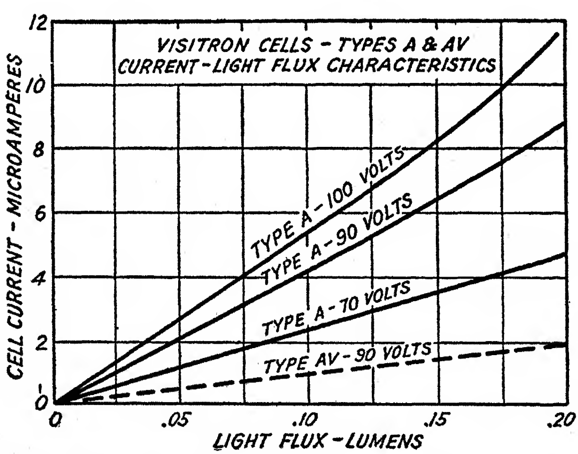


Fig. 500. Current-Light Flux Characteristic Curves of Photo-Cells. (Courtesy G-M Laboratories)

the photoelectric electrons ionize the gas on their way to the anode, and this ionization current is added to that of photo-

electric origin. At voltages above 90 with high illumination, the ionization becomes self-sustaining, which is detrimental to the tube.

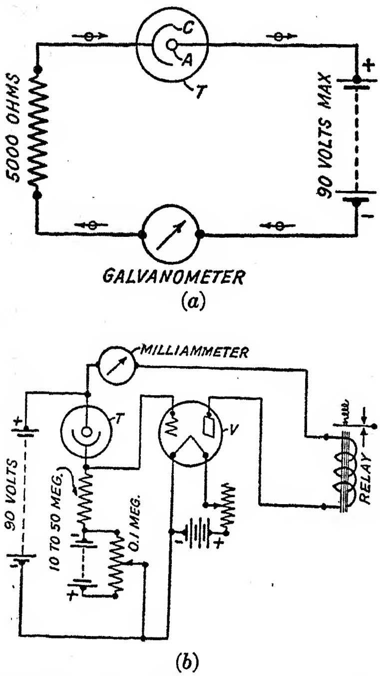


FIG. 501. Connections for Photo-Cell. (Courtesy G-M Laboratories)

In Fig. 500, photoelectric current is plotted against light flux (lumens), and it is seen that for the vacuum type the current is directly proportional to the intensity of the incident beam of light. But with the gas-filled type this is not true at the higher values of illumination.

The simplest connection, for such purposes as comparing light intensities, is shown in Fig. 501. Light falling on the sensitive surface of the cathode C causes the emission of electrons which are attracted to the anode A, the resulting current being indicated on the galvanometer.

The photoelectric cell, or, as it is often called, "the electric eye," has not the same color response as the human eye (Fig. 502). The color response is also different for tubes employing different sensitive substances. Nevertheless it has found a wide range of usefulness in everyday affairs. Among its many applications may be mentioned counting devices, machines for sorting on the basis of color, talking motion pictures, television, telephotography, photometry, and many others.

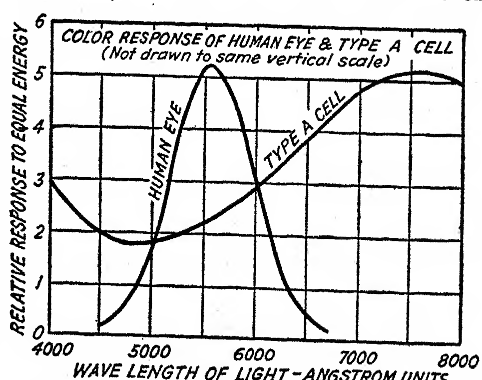


FIG. 502. Color Response of Photo-Cell and Human Eye. (Courtesy G-M Laboratories)

523. The photovoltaic effect. The *photovoltaic effect* is utilized in the *photronic cell* (Fig. 503) of the Weston Electrical Instrument Company. The light incident upon the cell develops an emf sufficient to produce a current of the order of 1.4 microamperes per foot-candle (120 microamperes per lumen) if the external resistance is small. The cell therefore requires no auxiliary battery. It should not be subjected to an external voltage.

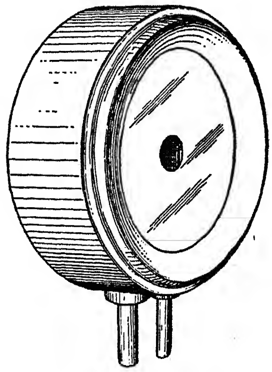


FIG. 503. Photronic Cell. (Courtesy Weston Electrical Instrument Co.)

The photronic cell "consists of a metal disc on which is a film of light-sensitive material.* The metal disc forms the positive terminal and a metal collector ring in contact with the light-sensitive surface forms the negative terminal." These elements are mounted in a bakelite case having a glass or quartz window, according as it is intended for visible or ultraviolet light. No vacuum is required.

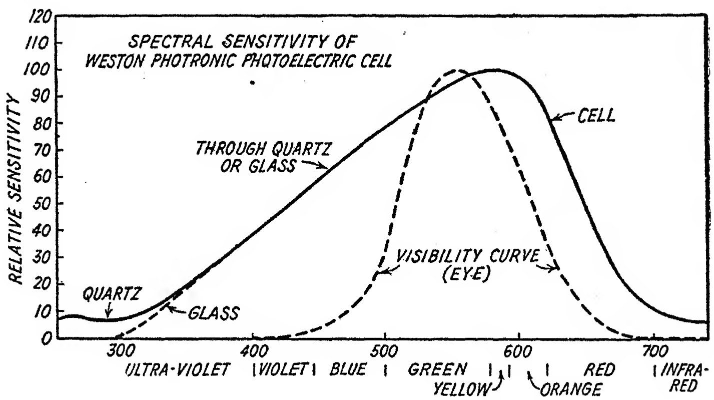


FIG. 504. Color Response of Photronic Cell and Human Eye

The action of the cell "appears to be entirely electronic, and all tests so far made indicate that no chemical or physical change takes place, so that the life of the cell seems unlimited."

The color response is much like that of the human eye, as will be seen from Fig. 504.

* The phenomenon is exhibited by cuprous oxide on copper.

CHAPTER XXXVIII

ELECTROMAGNETIC WAVES—RADIO

524. Electric oscillations. We have seen that in electricity, inductance L plays the role of inertia and capacitance C corresponds to the reciprocal of elasticity. From such considerations, Lord Kelvin predicted, in 1853, that, since a weight suspended on a spring executes oscillations, a current set up in a circuit which has inductance and capacitance should oscillate; i.e., it should flow back and forth. For the period T of an oscillation when the resistance is negligible, he deduced the value:

$$T = 2\pi\sqrt{LC} \tag{409}$$

where L is in abhenries and C is in abfarads, or
L is in henries and C is in farads; and
T is in seconds for both systems of units.

Consequently, since the frequency n is always the reciprocal of period,

$$n = \frac{1}{2\pi\sqrt{LC}} \tag{410}$$

where n is in cycles per second.

Lord Kelvin's prediction was confirmed in 1857 by Feddersen, who showed that an electric spark was oscillatory by observing it in a rotating mirror. These relations now form the basis of all our work in radio.

525. Electromagnetic waves. In the process of giving mathematical expression to Faraday's experimental facts about electricity and magnetism, James Clerk Maxwell achieved one of the greatest triumphs of mathematical physics. From his equations he predicted in 1864 that electric and magnetic fields produced by an oscillating current should travel outward into space in the form of electromagnetic waves with the speed of light, and should have the general nature of waves of light. Such waves were first

produced in 1888 by Heinrich Hertz, who found them to possess all the properties of waves of light, as anticipated by Maxwell. They pass readily through most substances except conductors: conductors reflect them.

Though of the same nature as light, these waves are too long to

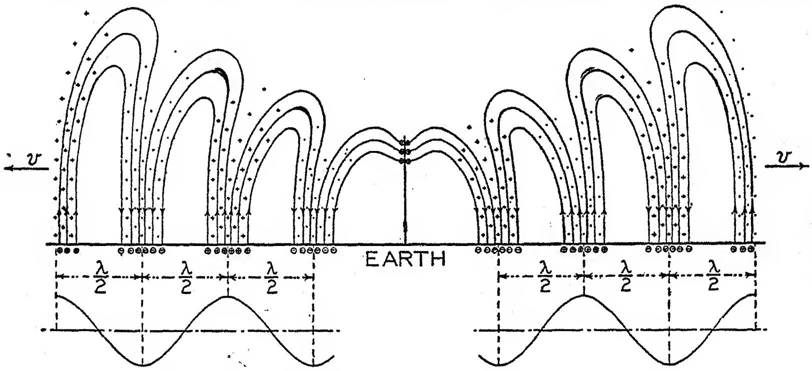


FIG. 505. Cross-Section of Electromagnetic Waves

be perceived by the eye; but they may be represented graphically as shown in Fig. 505. Here we have a vertical antenna in which an alternating current is maintained by any type of alternator having one side grounded. The antenna and the earth are then the two plates of a condenser, and alternating electrostatic and magnetic fields are produced about the antenna, the electrostatic lines being in vertical and the magnetic lines in horizontal planes. With reversed current in the antenna, both fields are reversed.

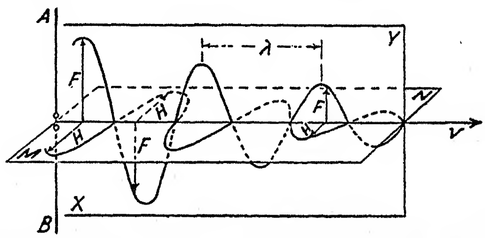


FIG. 506. Curves of Electric and Magnetic Intensity

The radiated waves are not material waves like sound waves in air, but consist of variations of electric and magnetic field intensity at the points in space through which they pass. Figure 506 indicates how the electric field intensity F and the magnetic field intensity H are always at right angles to each other and to the velocity of propagation v. It also shows how they vary sinusoidally, decreasing in amplitude as they proceed but maintaining the wave length λ unchanged.

526. Electric resonance. The phenomenon of resonance, i.e., the production of vibrations in one system by vibrations of the same frequency in another system, was demonstrated for electricity by Sir Oliver Lodge in 1888, just prior to the classic experiments of Hertz with electromagnetic waves.

Lodge's "resonant Leyden jar experiment," as it is often erroneously called, is made with two similar electric systems (Fig. 507), each consisting of a loop of wire (inductance) connected to the two coatings of a Leyden jar (capacitance) and a spark gap (indicator). Sparks are produced across the gap S_1 by connecting its sides to the high voltage terminals of an induction coil. These sparks are evidence that electricity is surging back and forth in

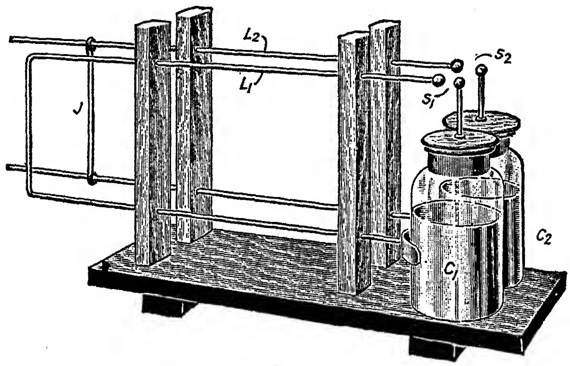


Fig. 507. Lodge's Resonant Leyden Jars. (Courtesy Central Scientific Co.)

the system L_1C_1. On account of this oscillating current, electromagnetic waves are sent out from the loop of system L_1C_1.

If the slider J is then carefully adjusted until the frequency of the second system L_2C_2 is the same as the frequency of the first system, sparks will spring across the gap S_2, indicating that some of the energy of the electromagnetic waves has been absorbed by the second system, resulting in an oscillatory current in that system.

Lodge's resonance experiment was wireless communication in embryo. Of the many scientists working on this problem, the first to develop a system giving transmission over a distance of several miles was Guglielmo Marconi of Bologna. On June 2, 1896, he received British Patent No. 12039 for his inventions.*

* J. A. Fleming, *Principles of Electric Wave Telegraphy* (New York, Longmans, Green and Co., 1910), p. 519.

527. The wavemeter. This simple instrument may equally well be called a frequency meter, since wave length and frequency are connected by the well-known relation of Sec. 216:

$$v = n\lambda.$$

Since the velocity of electromagnetic waves is the same as that of light (300,000 km/sec), and since, by Eq. (410),

$$n = \frac{1}{2\pi\sqrt{LC}}$$

it follows that

$$\lambda = \frac{v}{n} = \frac{3 \times 10^8}{n} = 3 \times 10^8 \times 2\pi\sqrt{LC}$$

$$= 1885\sqrt{LC} \times 10^6 \qquad (411)$$

where λ is the wave length in meters;
L is in henries; and
C is in farads.

This relation and the phenomenon of resonance give us the elements for constructing a wavemeter (Fig. 508). This consists of a circuit containing a coil of inductance L, a condenser of variable capacitance C, and an indicator G, which is often a thermo-galvanometer or a miniature lamp.

The coil of the wavemeter is placed so as to interlink with the magnetic field of the circuit under test, and the capacitance of the condenser is varied until the wavemeter circuit is in resonance with the other circuit. This condition is shown by maximum current in the indicator.

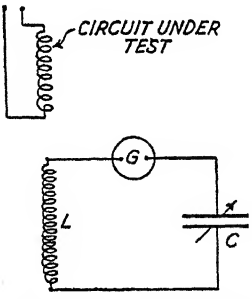


FIG. 508. Wavemeter Circuit

If both L and C are known, the wave length may be computed by Eq. (411); but it is customary to draw a calibration curve for the instrument, plotting wave lengths against dial readings of the condenser. For any setting of the dial, the wave length may then be read directly from the curve.

528. Transmitting systems. The earlier spark transmitting sets sent out a train of damped waves corresponding to each spark, and the frequency heard in the receiving phones was the number

of sparks per second. This system did not lend itself to telephony, and is now seldom used except in emergency sets.

After the ability of the triode to serve as an oscillator was understood, continuous waves of radio frequency (i.e., above 10,000 cycles per sec) came to be used for most purposes. In telegraphy they are broken up into groups that by their time length represent dots and dashes, this being accomplished by means of a key in either the grid or the plate circuit.

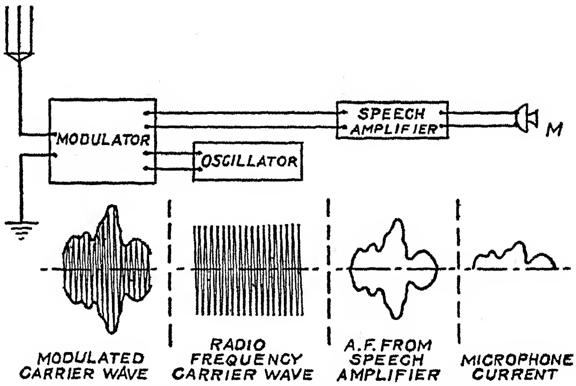


FIG. 509. Diagram of Broadcasting System

Broadcasting of speech and music is accomplished in practically all cases by some modification of the system outlined diagrammatically in Fig. 509. The sound waves impinging upon the disk of the microphone M produce variations in the current through the microphone as shown directly above it. This fluctuating "speech current" is amplified and converted into an audio-frequency alternating current in the speech amplifier. The alternating low-frequency current then passes to the modulator, where it is superimposed upon the radio-frequency carrier wave produced by the oscillator. The result is the modulated carrier wave which is radiated from the antenna.

A simple hookup for such a system is shown in Fig. 510, the amplifier unit being omitted.

529. Receiving systems. In all receiving sets, the antenna, or aerial, system absorbs some energy from the electromagnetic waves that encounter it, and an oscillating current is thereby set up in it. The larger the aerial and the more nearly it is in reso-

nance with the waves to be received, the greater will be the oscillating current set up. The function of the rest of the receiving set is to reproduce from this energy the sounds that originally gave form to the electromagnetic waves.

The connections for the triode as a detector (Sec. 517) give a

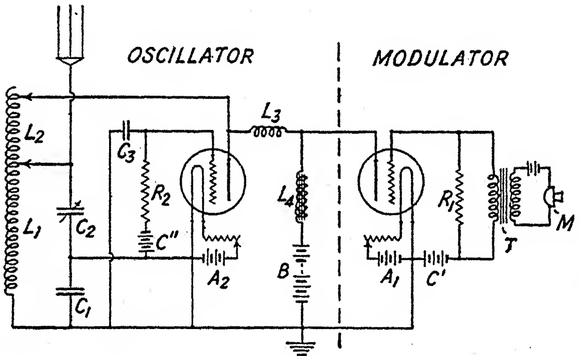


FIG. 510. Circuits of Broadcasting System

simple type of receiving set, which may be combined with one or more tubes connected as amplifiers (Sec. 518).

Detection changes the radio-frequency alternating currents into unidirectional currents that vary at audio-frequencies and hence are able to actuate the phones or loudspeaker. Amplification may be effected either before or after detection. In the former case it is called **radio-frequency amplification**; in the latter, **audio-frequency amplification.**

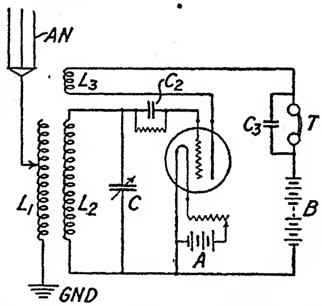


FIG. 511. Circuit of Receiving System

A **regenerative** type of receiver is shown in Fig. 511. Part of the energy taken from the electromagnetic waves by the antenna-to-ground system is passed by the loose-coupled, air-core transformer L_1L_2 to the circuit L_2C, which is tuned to the frequency of the waves by the variable condenser C. The variations of potential thus impressed upon the grid produce a partially rectified and amplified radio-frequency current in the plate circuit. (The method of rectification with the grid condenser C_2 and grid leak is somewhat different from the grid-bias method of Sec. 517.)

Some of the energy of the partially rectified current in the plate circuit is fed back into the grid circuit by the magnetic coupling between the "tickler" coil L_3 and the coil L_2, and is again amplified. This "feed-back" process is repeated indefinitely and greatly increases the strength of the signals in T. If carried too far, however, the set becomes an oscillator, sending out electromagnetic waves of its own, which, by heterodyning with other waves, produces a whistling interference in neighboring sets. Unfortunately this receiver is most effective just before it begins to oscillate.

530. The superheterodyne receiver. Most of the standard receiving sets today employ the heterodyning principle in one

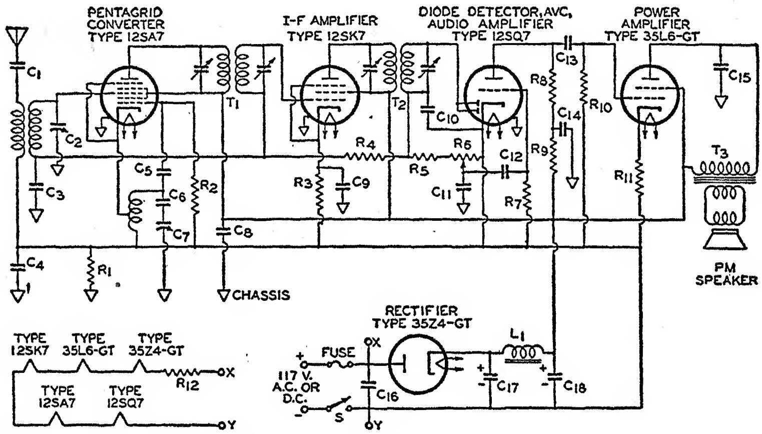


FIG. 512. AC–DC Superheterodyne Receiver. (Courtesy Radio Corporation of America)

form or another. The heterodyne method consists in superimposing upon the radio-frequency signal current another high frequency current generated in an oscillator which is part of the set. The mixing of these two radio frequencies produces *beats* (Sec. 234), the beat frequency being equal to the difference between the two component frequencies. This beat frequency is called the "intermediate frequency," and it is at this frequency that the current is greatly amplified, then again detected, and further amplified to give the final speech reproduction.

The hookup of an a-c operated superheterodyne receiver is shown in Fig. 512. Here the antenna current passes first to a stage of radio-frequency amplification *R.F.*, and then to the first de-

tector, where it is combined with the radio-frequency current from the local oscillator *OSC*. This produces the intermediate, or beat, frequency to which the intermediate amplifier *I.F.* is tuned. From this amplifier it goes to the second detector, thence to the power amplifier, and finally to the loudspeaker. The arrangement illustrated is representative of the many possible connections for this system of reception.

The superheterodyne receiver has certain advantages over other types of receiver. Its selectivity is usually greater, because two stations broadcasting at 1000 kc/sec and 1010 kc/sec differ by only 10 kc/sec, which is a difference of 1%. After heterodyning to an intermediate frequency of, say, 175 kc/sec, they still differ by 10 kc/sec, which is about a 10% difference, so that now the stations are more easily separated. Another advantage is that most of the amplification is secured at the comparatively low intermediate frequency (about 175 kc/sec) at which each intermediate stage can be caused to give an amplification of 60 or 80, if screen-grid tubes are used; whereas an amplification of, say, 40 is about all that can be obtained at high frequencies (1500 to 550 kc/sec) without disagreeable effects due to stray fields and tube capacitances. Consequently, the superheterodyne requires very little energy to be absorbed from the electromagnetic waves in the ether, and hence needs only a very small aerial.

It is considerations such as these that have given to this type of receiver its well-deserved popularity.*

530a. Television. Just as the telephone translates the energy of sound waves into that of electric currents, and retranslates these currents back into sound waves again at the receiver, so the television transmitter translates the energy of light waves into that of radio waves which are retranslated into light waves at the receiving end of the system.

The transmitter consists of a vacuum tube called an **iconoscope,** Fig. 512a. The picture to be transmitted is focused by the lens *L* upon the mosaic side of the target *T*. This target consists of a

* References on Radio:
Ghirardi, *Radio Physics Course* (New York, Radio Technical Publishing Co.).
Terman, *Radio Engineering* (New York, McGraw-Hill Book Co.).
Radio Amateur's Handbook (West Hartford, Conn., American Radio Relay League, Inc.).

sheet of mica about 9 cm by 12 cm by 0.005 cm thick, the back side of which is coated with a continuous metallic film. But the side toward the lens is a mosaic consisting of an enormous number of microscopic caesiated silver particles approximately 0.0005 cm in diameter. Each of these particles with the metallic film on the back of the mica forms a tiny condenser.

When the picture to be transmitted is focused on the mosaic, each caesiated particle emits a number of photoelectrons depending upon the intensity of the light at that point, and thus becomes +. The electrons pass to the collector ring shown dotted.

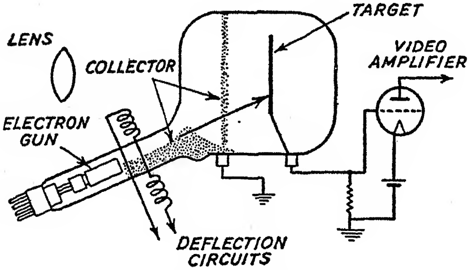


Fig. 512a. Iconoscope. (Courtesy of Professor Knox McIlwain, *Jour. Appld. Phys.*, July, 1939)

By means of magnetic fields from external coils, an electron beam 0.02 cm in diameter is caused to sweep across, or "scan," the mosaic in 441 horizontal strips, the whole picture being thus scanned 30 times per second.

When the scanning beam strikes a particle of the mosaic, additional secondary electrons are emitted by the particle, the change of charge on the tiny condenser depending upon its previous charge which in turn depended upon its illumination. This change of charge on the little condenser produces a corresponding change of potential of the grid of the video amplifier. Thus a radio signal is sent out whose energy depends upon the illumination of the mosaic particle under consideration.

The television receiver or **kinescope** (Fig. 512b) is a cathode ray tube quite similar to that of Sec. 506. The electron source is an internally heated cylindrical cathode having a disk-shaped cap coated with barium and strontium oxides.

The video signal picked up by the antenna is impressed upon the control grid and determines the number of electrons in the beam. Electrostatic fields, serving as electron lenses (Sec. 623-1), direct these electrons to a fine focus on the internal surface of the end of the tube which is coated with a fluorescent material, or phosphor.

By means of magnetic fields from external coils, the electron beam is made to execute a scanning motion identical to and in synchronism with that of the scanning beam of the iconoscope. Synchronism is maintained by signals radiated from the sending station along with the video signals.

Hence when the scanning beam in the iconoscope passes over a mosaic particle, a signal is sent out which causes the reproducing

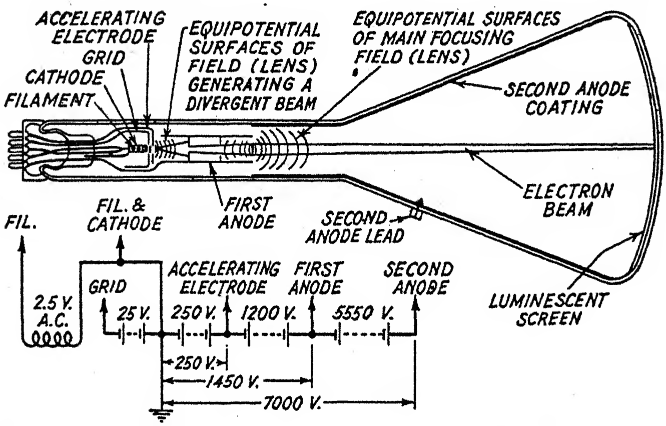


FIG. 512b. Television Receiver. (Courtesy of Dr. E. W. Engstrom, *Jour. Appld. Phys.*, July, 1939)

beam of electrons to strike a similarly located spot on the luminescent screen of the receiver; and this spot fluoresces with an intensity proportional to the illumination on the mosaic particle. Consequently the picture projected on to the mosaic screen of the iconoscope is reproduced on the fluorescent screen of the receiving cathode ray tube.

LIGHT

Accident favors only those who are prepared.
—Pasteur

CHAPTER XXXIX

NATURE AND PROPAGATION

531. Historical. Optics, or the science of light, is one of the oldest branches of physics. A converging lens of quartz was found in the ruins of Nineveh (destroyed 606 B.C.), and the use of such lenses to kindle fires is mentioned by Aristophanes (424 B.C.).

The followers of Pythagoras (500 B.C.) thought that vision resulted when particles projected from an object entered the eye, while the school of Plato (400 B.C.) held the fantastic doctrine that vision ensued when "the divine fire" emitted by the eye mixed with an emanation from the body seen.

The **corpuscular theory of Sir Isaac Newton** (about 1666), like that of Pythagoras, considered light to consist of corpuscles, or particles, shot out by the observed body. This theory accounted satisfactorily for the fact that light travels in straight lines and for reflection, but it led to the conclusion that the speed of light in water should be *greater* than in air.

The **wave theory of Huyghens** (1678) was the rival of that of Newton for a hundred years. On the basis of Huyghens' principle, it explained reflection and refraction, but it predicted that the speed of light in water would be *less* than in air.

On account of Newton's great prestige, his theory prevailed until the speed of light in water was actually measured by Foucault in 1850 and found to be about three-fourths of the speed in air. The corpuscular theory then died temporarily.

Huyghens and his supporters considered light waves to be mechanical waves, longitudinal like those of sound; but on this basis it was not possible to account for the phenomenon of polarization. Young and Fresnel observed independently in 1817 that polarization could be explained if the waves were transverse. The physical nature of such waves remained obscure, however, until the time of James Clerk Maxwell.

The **electromagnetic theory** of light advanced by Maxwell in

1873 identified light waves with the transverse electromagnetic waves whose existence he predicted from mathematical reasoning and which were subsequently discovered by Hertz in 1888. These are not mechanical waves, but are waves of progressively changing electric and magnetic field intensities. We are accustomed to think of waves as existing in some medium; yet light traverses interstellar space, which is practically devoid of matter. This philosophical difficulty was overcome by hypothesizing the presence of a medium called "the ether," which was assumed to fill all space—even that occupied by material bodies. The actual existence of this medium is still a moot question.

But wave theories failed to explain spectra.

The **quantum theory of Max Planck** was applied by Niels Bohr in 1913 to the atom model suggested by Rutherford, with the result that a theory was evolved which, though not consistent with the wave theory, offers rational explanations of many spectral phenomena.

At present, therefore, from the physical standpoint, the evidence indicates the following definition:

Light is radiant energy capable of producing the sensation of vision. It is radiated in quanta $h\nu$ and transmitted through space in electromagnetic waves.

The factor h is Planck's universal constant, and ν is the frequency of the light waves under consideration.

The new **quantum mechanics of Heisenberg** (1925) and the concept of "material waves" suggested by Louis de Broglie in 1924 give promise of yielding a single consistent theory of the nature of light and of matter.

532. Rectilinear propagation. It is a matter of common observation that, when sunlight enters a somewhat darkened and dusty room through a keyhole or other small aperture, the beam of light appears to be perfectly straight. An engineer can run a straight road because the light from the tacks in the tops of his stakes comes to his eye along straight lines.

On further investigation, however, we find that when the density of the air or other medium is not uniform, as in the case of a mirage (Fig. 513) or above a hot stove, the beam of light is bent. Hence we must say:

Light travels in straight lines only in homogeneous media, i.e., in media which have the same properties in all parts.

Moreover, from his general theory of relativity, Einstein predicted that a beam of light would be bent while passing through a gravitational field. Subsequent verification for light passing through the intense gravitational field near the sun is an element of evidence of the correctness of the relativity theory. Even near the sun, this gravitational deflection is very slight (only 1.7 sec of arc); hence it is negligible except in refined astronomical work.

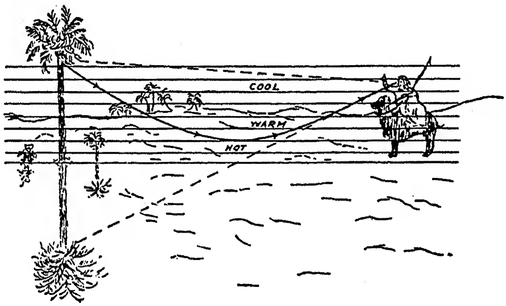


FIG. 513. A Mirage

From a luminous particle, light is radiated equally in all directions in spherical waves having their common center at the particle.

A **ray** is any radius of this system of spherical waves, and merely indicates the direction of propagation.

A **beam,** or pencil, of light is the portion of a system of spherical waves comprised within a pyramid having its vertex at the common center of the spheres. If the center is very remote, so that the wave fronts are sensibly parallel, the beam is called a **parallel beam.**

533. Shadows. A shadow is the region from which light is cut off, wholly or partially, by an intervening body.

If the light is from a point source P, the **geometrical shadow** is limited by lines drawn from the source to each point on the contour of the intervening body B, as shown in Fig. 514. When intercepted by a screen S, a geometrical shadow is bounded by a perfectly definite, or sharp,

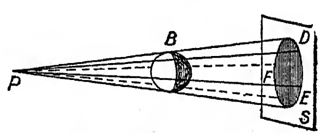


FIG. 514. Shadow—Light from Point Source

line DEF. Actual shadows, however, never have this sharp boundary. They are slightly hazy at the edge because we never have strictly a point source, and also because of diffraction (Sec. 222).

If the source is not a point, the shadow consists of two parts: the **umbra,** or the region from which all light is cut off; and the **penumbra,** or partial shadow, from which a part of the light is cut off—varying from all at the edge of the umbra to none at the outer edge of the penumbra. As shown in Fig. 515, the umbra is defined by the external tangents AB and CD, and the penumbra by the internal tangents AD and CB.

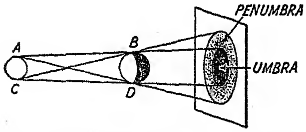


FIG. 515. Shadow—Light from Large Source

534. Pin-hole images. If light from a luminous body B (Fig. 516) is permitted to pass through a small hole in a screen S_1 and is intercepted by a second screen S_2, an inverted image of the object is produced on S_2.

Each point of the luminous body is the vertex of a cone of waves determined by the edge of the hole in S_1. The closer S_2 is to S_1, the larger will be the bases of the cones relative to the size of the image as a whole, and hence the less distinct will be the image on account of the overlapping of these bases.

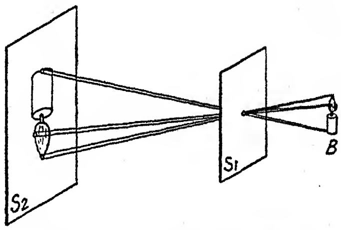


FIG. 516. Pin-hole Image

The fact that the image is inverted is seen to be a necessary consequence of the rectilinear propagation of light.

535. The speed of light: Römer's method. The first recorded attempt to measure the speed of light was made by Galileo. His method was to station two observers A and B at night on two hilltops about a mile apart, each provided with a dark lantern. A uncovers his light. On seeing A's light, B uncovers his. A observes the time elapsing from the instant of exposing his light until he sees the flash from B's lamp. The net result was that *no* time appeared to be required for light to travel this distance.

More successful was the Danish astronomer, Olaf Römer, working at Paris where he had been invited to reside by Colbert, director-general of Louis XIV. While observing the eclipses of

one of Jupiter's moons, he noted in 1676 that the eclipse occurred approximately 500 sec ahead of the predicted time when the earth was at E_1 (Fig. 517), and the same amount too late when the earth was at E_2.

He concluded correctly that this difference of nearly 1000 sec was the time required for light to traverse the diameter of the earth's orbit (186,000,000 mi).

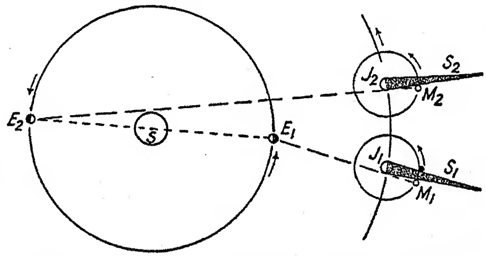


FIG. 517. Speed of Light—Römer's Method

This gives 186,000 mi/sec as the speed of light. On account of an error in his value of the orbit's diameter, Römer obtained 192,000 mi/sec for the velocity, which was too high. But it was an excellent determination for that time and remained the only one until Bradley, in 1727, secured a better one from the aberration of light.

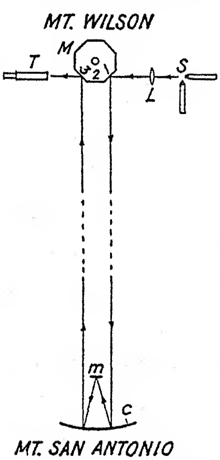


FIG. 518. Speed of Light — Michelson's Method

536. The speed of light: Michelson's method. In the years 1880–82, A. A. Michelson made very precise determinations of the speed of light at the Case School of Applied Science in Cleveland, Ohio, obtaining the value 299,853 km/sec. This was the accepted value for half a century. He used the revolving mirror method devised by Foucault in 1850, with improvements. In 1926–29, he repeated the experiment using the same method but with great refinement of apparatus and technique. The method is shown in principle in Fig. 518.

In the later experiment, the revolving octagonal mirror M was located on Mount Wilson and the stationary mirror C on Mount San Antonio in California. The distance between these stations, measured by the U. S. Coast and Geodetic Survey, was 35,426.3 ± 0.1 meters (about 22.5 mi).

The apparatus was so adjusted that when M was at rest, a narrow beam of light from the arc lamp S was reflected from the polished face 1 of M to the concave mirror C, and thence to the

plane mirror m. The latter reversed the direction of the beam, so that it was returned by C to face 3 of M, whence it was reflected onto the cross hairs of the telescope T.

The mirror M was then set in rotation clockwise and its speed, which could be controlled and measured with extreme accuracy, was increased until face 2 caught the beam of light returning from C and reflected it again upon the cross hairs of T. This meant that in the time required for light to make the round trip (70.8526 km), the mirror made 1/8 of a revolution.

The rotational speed was then observed to be 528.76 rev/sec. Hence the light traveled the above distance in 1/8 of 1/528.76, or 1/4230.08 sec. The speed of the light was therefore 70.8526 divided by 1/4230.08, which gives the value 299,712 km/sec *in air*.

When reduced to what it would have been if the experiment had been performed *in a vacuum* the speed of light is 299,796 $\pm$ 4 km/sec, or 186,285 mi/sec, which is now the accepted value.

The most recent experiment (1931–33) was made at Santa Ana, California, by Michelson, Pease, and Pearson. The same general method described above was followed. Using a rotating mirror of 32 faces and turning 730 rev/sec, the beam of light was reflected back and forth in a steel tube 1 mi long and 36 in. in diameter until optical paths 8 and 10 mi respectively were secured. The tube was exhausted to a pressure of from 0.5 to 5 mm of Hg. The work was completed by Pease and Pearson after the death of Professor Michelson in 1931, and the results were published in 1934.

The mean of 2885 determinations gave 299,774 km/sec as the speed of light in vacuo, the average deviation being 11 km/sec.

537. Speed of light in other media. We usually find the speed of light in a material medium from its speed in air and its index of refraction (Sec. 221); but both Foucault and Michelson made actual measurements of its velocity in water by placing tubes of water in the path of the beam of light of their apparatus. In this way the speed of light in water was found to be about three-fourths of its speed in air.

As previously mentioned, this fact overthrew Newton's corpuscular theory, which required that the speed in water should be greater than in air.

538. Dependence of speed on color. In material media such as glass and water, the speed of red light (long waves) is greater than that of blue light (short waves), as is shown by the fact that the index of refraction is usually less for red than for blue light.

If this were true in a vacuum also, one of Jupiter's moons would appear blue just as it disappears behind the planet, and red just as it emerges. But it does not do that: no such color change has ever been observed. The space between Jupiter and the earth is practically a vacuum.

Hence we conclude that **lights of all colors (frequencies) travel with the same speed in a vacuum.**

PROBLEMS

1. If the distance from the earth to the sun is 93,000,000 mi, how long does it take light to come from the sun to the earth?

2. Before watches were made, Galileo attempted to measure the time for light to travel 2 mi. Why did he fail? Explain arithmetically.

3. The nearest star, Alpha Centauri, is 4.3 light-years distant. How far is that in miles? (A light-year is the distance light travels in one year.)

4. The speed of radio waves is the same as that of light. How long will it take a radio signal to go around the earth?

5. What is the wave length of radio waves whose frequency is 500 kc?

6. In a New York church, one auditor is 100 ft from the speaker while another in Miami hears him by radio. Which auditor will hear him first? (Distance from New York to Miami = 1200 mi.)

7. The wave length of sodium light (yellow) is 0.00005893 cm. How many of these are there in 1 cm? What is the frequency?

8. When light passes from one medium into another the frequency remains unchanged. If the speed of sodium light in water is three-fourths of its speed in air, what is the wave length of sodium light in water?

9. If the total path length between reflections at the rotating mirror in Michelson's method of determining the velocity of light was 50 mi, what would be the least rotational speed of a 32-sided mirror to bring the narrow beam of light back to the cross-hair?

CHAPTER XL

REFLECTION—MIRRORS

539. Law of reflection. **The angle of incidence *i* and the angle of reflection *r* are equal and lie in the same plane** (Fig. 176). This law has already been demonstrated for all types of waves by Huyghens' principle in Sec. 219; hence it is true also for light.

540. Regular and diffuse reflection. A **mirror** is a body whose surface is highly polished so that distinct images of other bodies

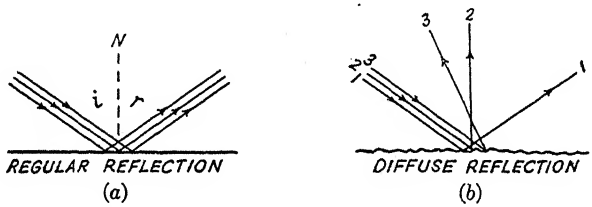


FIG. 519. Regular and Diffuse Reflection

may be seen in it by reflection. If it is a plane surface, rays that are parallel before reflection are still parallel after reflection; if its surface is curved, rays are still reflected in a perfectly definite, or regular, way. Such reflection is called **regular reflection** (Fig. 519a).

But when light falls upon a matte, or slightly rough, surface such as plaster or blotting paper, the rays are scattered in all directions. Each individual ray obeys the law of reflection at the tiny portion of the surface on which it strikes; but rays that were parallel before the reflection are turned in random directions, and are not parallel after reflection. Such reflection is called **diffuse reflection** (Fig. 519b).

All non-luminous bodies are seen by the light that they reflect diffusely. A perfect mirror surface could not be seen; but actual mirrors usually reflect some light diffusely on account of slight irregularities in their surfaces, and therefore can be seen.

541. Plane mirrors. In Fig. 520, let AOB and APC be two paths of light from any point A of an object in front of a plane mirror MM'. Since light enters the eye from the directions OB and PC, it appears to come from the point A', which is called the **image** of A.

From the geometry of the figure, it is seen that the image A' of point A is as far behind the mirror as A is in front of it. If we make the construction for two points of an object, we find that the distance between their images is the same as the distance between the points.

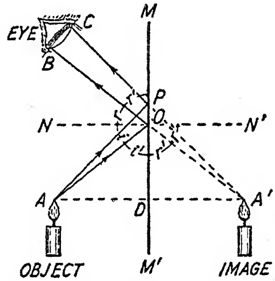


FIG. 520. Image in Plane Mirror

Therefore, **the image of an object appears as far behind a plane mirror as the object is in front of the mirror; and the size of the image is the same as the size of the object.**

542. Spherical mirrors. Next to plane mirrors, spherical mirrors (Fig. 521) are the type most commonly used.

A **concave** spherical mirror is made by polishing the **inner** surface of a segment of a sphere.

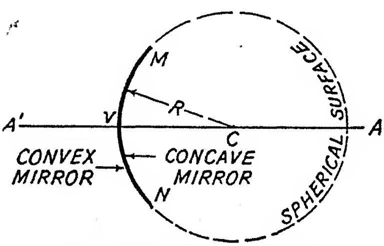


FIG. 521. Spherical Mirrors

A **convex** spherical mirror is made by polishing the **outer** surface of a segment of a sphere.

The following terms apply to spherical mirrors:

The **center of curvature** C of the mirror is the center of the sphere.

The **radius of curvature** R of the mirror is the radius of the sphere. For a convex mirror the radius is considered negative.

The **vertex** V is the pole, or central point, of the segment MVN that forms the mirror.

The **principal axis** AA' is the line joining the center C to the vertex V.

A **secondary axis** is any other straight line through the center.

The **aperture** MN of the mirror is the diameter of the small circle which is the base of the spherical segment.

543. Principal focus. The principal focus F of a spherical mirror is the point on the principal axis through which all rays that were parallel to the principal axis before reflection, pass after reflection.

For spherical mirrors whose aperture is small compared to the radius, the principal focus F lies approximately halfway between the center of curvature C and the vertex V.

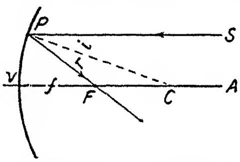


FIG. 522. Concave Mirror

This may readily be shown geometrically, as in Fig. 522. Let SP be **any** ray parallel to the principal axis AV of a spherical mirror whose center is at C. The radius CP is the **normal** at P; hence PF is the reflected ray, making

$$\angle r = \angle i \quad \text{(by the law of reflection).}$$

In $\triangle PFC$.

$$\angle PCF = \angle i \quad \text{(alternate interior angles).}$$

Therefore $$\angle PCF = \angle r \quad \text{(each being equal to } \angle i)$$

and $$PF = FC \quad \text{(sides opposite equal angles).}$$

But if arc PV is small compared to the radius R

$$PF \doteq VF$$

and hence, $$VF \doteq FC.$$

That is, the focus lies halfway between the vertex and the center of curvature.

The distance f from the vertex to the focus is called the **focal length** of the mirror. From the foregoing, it follows that

$$\text{Focal length, } f \doteq \frac{R}{2}. \tag{412}$$

By similar reasoning, the same relations may be shown to be true for convex spherical mirrors; but both f and R lie behind the mirror and hence are considered **negative**.

544. Principal focus: Method of waves. Since light is transmitted as electromagnetic waves, the method of waves may be considered the logical way of solving its problems. As an illus-

tration of this method, the relation of the preceding paragraph will be demonstrated for a convex mirror.

Let a parallel beam of light moving from right to left, as shown by the ray SO (Fig. 523a), impinge upon the convex mirror MVN. If the mirror were not present, the wave front would continue its motion undisturbed to the dotted position MQN. But when the center of the wave front strikes the mirror at V, it is turned back; and by Huyghens' principle the wavelet sent out from V reaches the point P in the same time (one period) in which it would have reached Q without the mirror. By constructing the wavelets from other points of the mirror surface as shown, we obtain the reflected wave front OPN, whose center is at F.

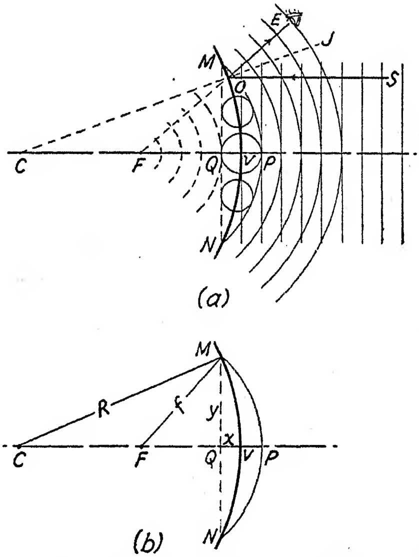


FIG. 523. Location of Focus by Method of Waves

The light that enters the eye E, as shown by the reflected ray OE, appears to originate at the point F behind the mirror, since F is the center of the reflected wave system that reaches the eye. A focus is a point from which light diverges, or to which it converges. Consequently, F is the principal focus of the mirror, since the waves were parallel before reflection.

In order to avoid confusion in Fig. 523a, the parts of the figure necessary for determining the location of F have been redrawn in Fig. 523b. There it will be seen that

$$y^2 = R^2 - (R - x)^2 = R^2 - R^2 + 2Rx - x^2$$

$$x = \frac{y^2}{2R - x}.$$

If the aperture angle of the mirror is not more than 5°, x will be small as compared to R and may be neglected, so that

$$x \doteq \frac{y^2}{2R}. \tag{a}$$

Similarly,

$$PQ = 2x \doteq \frac{y^2}{2f}. \tag{b}$$

Dividing Eq. (a) by Eq. (b),

$$\frac{1}{2} \doteq \frac{f}{R}.$$

Hence,

$$f \doteq \frac{R}{2}$$

which is the same as was found in Eq. (412) by the method of rays.

By putting this in the form:

$$\frac{1}{f} = 2\left(\frac{1}{R}\right)$$

we see that, since the curvature of the incident waves was zero, the mirror has impressed upon the waves a curvature $(1/f)$ which is twice the curvature $1/R$ of the mirror.

Many of the phenomena of light may be treated either by the method of waves (**physical optics**) or by the method of rays (**geometrical optics**). In most cases, however, the latter method gives more satisfying proofs by elementary geometry, and hence will be largely employed in the following pages.

545. Location of image formed by spherical mirror: Graphical method. The following procedure may be used to locate the image formed by either a concave or a convex spherical mirror (Fig. 524).

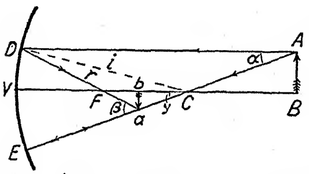


FIG. 524. Location of Image by Graphical Method

1. From any point A of the object draw a ray AD parallel to the principal axis. After reflection at D, this ray will pass through the principal focus F.

2. From A draw also a ray AE through the center of curvature C. This ray, being radial, is normal to the mirror; hence the light will be reflected back along the same path EA.

3. The point a where the reflected rays DF and EC meet is the *image of the point A*. All other rays from A that are reflected by the mirror will pass through a, except certain marginal rays that will be discussed later.

The points A and a are also called "conjugate foci" of the mirror. **Conjugate foci** are two points so situated that the light radiated from either is concentrated at the other. That is, conjugate foci are interchangeable: if a were the luminous point, its image would be A. A mirror will obviously have an infinite number of conjugate foci, since every point of the object has a conjugate focus in the image, and vice versa.

To simplify the mathematical treatment, let us call the

$$\begin{aligned}\text{Object distance} \quad AE &\equiv p\\ \text{Image distance} \quad aE &\equiv q\\ \text{Arc of mirror} \quad DVE &\equiv z.\end{aligned}$$

Then it is easily proven that

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}.$$

Proof. If the aperture of the mirror is small compared to the radius, we may write without serious error:

$$\angle\alpha = \frac{\text{arc } DVE}{AE} = \frac{z}{p} \quad \text{(by the definition of angle)}$$

$$\angle\beta = \frac{z}{q}$$

$$\angle\gamma = \frac{z}{R} \quad (\text{where } \gamma \equiv \text{angle } DCE).$$

Then

$$\alpha = \gamma - i \quad (\text{ext. } \angle = \text{sum of op. int. } \angle\text{s})$$
$$\beta = \gamma + r \quad (\text{ext. } \angle = \text{sum of op. int. } \angle\text{s})$$

whence, since

$$i = r \qquad \text{(by the law of reflection)},$$

we have

$$\alpha + \beta = 2\gamma$$

which, on substituting the above values, becomes:

$$\frac{z}{p} + \frac{z}{q} = \frac{2z}{R}.$$

Thus, **for a concave mirror,**

$$\frac{1}{p} + \frac{1}{q} = \frac{2}{R} = \frac{1}{f}. \tag{413}$$

If we carry through the same analysis, we obtain the relation, **for a convex mirror,**

$$\frac{1}{p} - \frac{1}{q} = -\frac{2}{R} = -\frac{1}{f}. \tag{414}$$

In deriving these equations, we have taken all quantities positive, just as they stand in the figures. By inspection of Eqs. (413) and (414) and the corresponding figures, it will be seen that Eq. (414) becomes the same as Eq. (413) if we adopt the following **convention of signs:**

1. f is to be taken + for concave and − for convex mirrors.
2. p is to be taken + when the object is real, and − when the object is virtual.
3. q is to be taken + when the image is real and − when the image is virtual.

When this convention is followed, the formulas for concave and convex mirrors are the same, and we have the **general mirror formula:**

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}. \tag{415}$$

546. Relative sizes of object and image. In Fig. 525, the image ab of the object AB has been constructed as described in the preceding section.

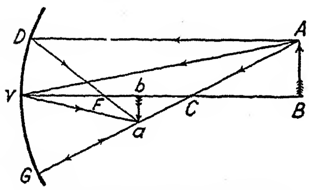


FIG. 525. Relative Size of Object and Image

Draw the ray AV. After reflection, it will pass through a, since A and a are conjugate foci. Also,

$\angle AVB = \angle aVb$ (by the law of reflection).

Hence,

$\triangle AVB$ is similar to $\triangle aVb$ (rt. $\triangle$s having an acute angle of one equal to an acute angle of the other).

Therefore,

$$\frac{ab}{AB} = \frac{bV}{BV} \quad \text{(homologous parts of similar figures).}$$

Or, stated in words,

$$\textbf{Magnification} \equiv \frac{\textbf{Size of Image}}{\textbf{Size of Object}} = \frac{\textbf{Image Distance from Mirror}}{\textbf{Object Distance from Mirror}}. \tag{416}$$

547. Real and virtual images. When waves of light actually converge to a point, i.e., when the rays used in determining an

image meet in an actual point a (Fig. 525), that point is said to be a **real image** of the point A from which the light emanates.

But if the waves only **appear** to converge to a point, i.e., the rays do *not* actually meet, as in Figs. 526e and 526f, the point a is said to be the **virtual image** of A.

The six cases that may occur are shown in Figs. 526a to 526f.

With concave mirror:

1. When the object AB is at a distance from the vertex V greater than the radius of the mirror, the image ab is between the center and the focus. It is real, inverted, and smaller than the object (Fig. 526a).

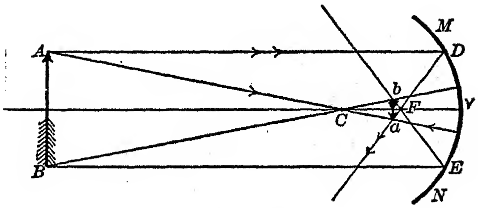


Fig. 526a

2. When the object is at the center of curvature, the image is also at the center of curvature. It is real, inverted, and the same size as the object (Fig. 526b).

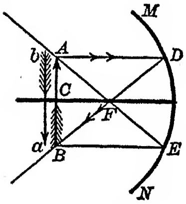


Fig. 526b

3. When the object is between the center and the focus, the image is beyond the center. It is real, inverted, and larger than the object (Fig. 526c).

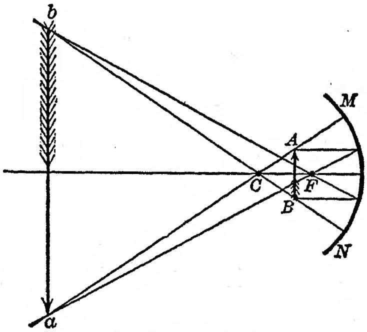


Fig. 526c

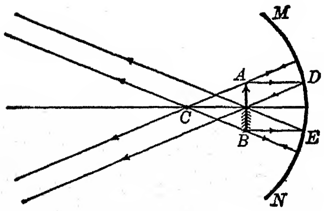


Fig. 526d

4. When the object is at the principal focus, the reflected rays are parallel. The image is at infinity; i.e., no image is formed (Fig. 526d).

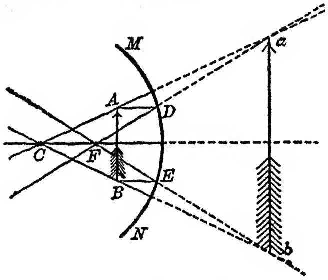


FIG. 526e

5. When the object is between the principal focus and the vertex, the reflected rays do **not actually** meet. They appear to meet behind the mirror. Hence the image is virtual, erect, and enlarged (Fig. 526e).

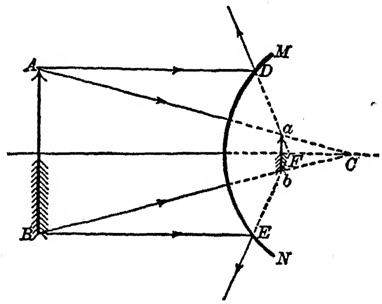


FIG. 526f

With convex mirror:

1. For all positions of the object, the image is virtual, erect, and smaller than the object (Fig. 526f).

These facts may be verified algebraically by substituting proper values of p and f in Eqs. (413) and (414). Thus, when $p = f$, we have:

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f} = \frac{1}{p}.$$

Therefore,

$$\frac{1}{q} = 0$$

and

$$q = \infty, \text{ as in Fig. 526d.}$$

From the above illustrations it will also be seen that, when only one mirror is used:

A **real** image is always **inverted,** and it may be caught on a screen.

A **virtual** image is always **erect,** and it cannot be caught on a screen.

548. Spherical aberration. Caustic by reflection. In deriving the equation for spherical mirrors, it was assumed that the angular aperture of the mirror was small—not exceeding 5°.

The reason for this may be understood from Fig. 527, where it

will be seen that, rays incident upon a spherical mirror from one point are **not all** brought to focus at another **single** point. This phenomenon is called **spherical aberration.**

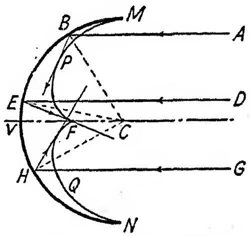


Fig. 527. Caustic by Reflection

Rays such as AB and GH, incident at points for which the angular distance from the vertex is large, do not pass through the focus F but envelop a curve $MPFQN$, known as the **caustic by reflection.** It may be seen as a bright curve on the surface of milk in a glass filled to within about an inch of the top, or within a napkin ring, when either is illuminated obliquely by a strong source.

On account of spherical aberration, images formed by spherical mirrors are always slightly blurred, instead of being perfectly clear and sharply defined. Better definition is secured by using a card with a hole in it to cut down the aperture of the mirror so as to use only those rays that are incident near the vertex.

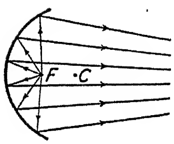


Fig. 528. Spherical Mirror Does Not Produce Parallel Beam

Another effect of spherical aberration is shown when a point source of light is placed at the principal focus of a concave spherical mirror. Then we do **not** secure a parallel beam of light, but a beam which at first converges (Fig. 528), and afterward diverges.

549. Parabolic mirror. The difficulty of producing a parallel beam by means of a spherical mirror is overcome by using a *parabolic mirror.*

The proposition will be recalled from analytic geometry that a normal AN to a parabola at a point A (Fig. 529) bisects the angle between the focal radius FA and a line through A parallel to the axis.

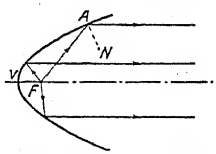


Fig. 529. A Parabolic Mirror Produces a Parallel Beam

Hence, if a mirror surface has the form of a paraboloid of revolution, and a point source is placed at the focus F, the reflected rays will be parallel to the axis. Parabolic mirrors are used for searchlights, the headlights of locomotives and automobiles, and for similar purposes requiring an intense

parallel beam. The great anti-aircraft searchlights project a parallel beam of 800,000,000 candlepower.

PROBLEMS

1. An arrow 2 in. long is placed with one of its ends 1 in. and the other end 1/2 in. from a plane mirror. Locate the image of the arrow and trace the reflected rays to an eye position taken at random.

2. Two plane mirrors are at right angles to each other. A pin is placed 2 in. from one mirror and 1 in. from the other. Locate all the images in both mirrors. Assuming the eye to be 4 in. from each mirror, trace the rays for each image to the eye.

3. Prove that the shortest plane mirror in which a person may see his entire length is one-half his height, and show how the mirror must be placed.

4. Show that if a plane mirror is rotated, the reflected beam turns through twice the angle through which the mirror turns.

5. An object 18 cm high stands 25 cm in front of a concave mirror of 40-cm radius. Find the position and size of the image.

6. An object 12 cm high stands 10 cm in front of a concave mirror whose radius of curvature is 60 cm. Find the position and size of the image.

7. An object is 70 cm from a concave mirror and the real image is 120 cm from the mirror. What is the radius of curvature of the mirror?

8. An object 40 in. in front of a concave mirror has its image 60 in. in front of the mirror. What is the radius of curvature of the mirror? If the object is 2 in. high, how high is the image?

9. If an electric lamp is 4 in. high and is 5 in. in front of a concave mirror of 12-in. radius, where is the image formed and how high is it?

10. When an object is 20 cm in front of a concave mirror, the virtual image is three times the size of the object. Where is the image located, and what is the radius of the mirror?

11. The moon is approximately 2000 mi in diameter and 240,000 mi away. What will be the size of its image formed by a concave mirror of 6-ft radius?

12. A concave mirror has a radius of 100 cm. Where should an object be placed to give a virtual image five times as large as the object?

13. Find the position and size of the image if the mirror in problem 9 were convex instead of concave.

14. A convex mirror has a radius of 24 in. If an object is 36 in. in front of the mirror, where is the image? If the object is 10 in. high, how high is the image?

15. If a convex mirror gives an image one-third the size of the object when the object is 30 cm in front of the mirror, what is its radius?

16. In order to find the focal length of a convex mirror, a candle is placed 4 cm in front of it and the light is reflected into a concave mirror of 16-cm radius, which forms a real image 24 cm in front. If the two mirrors are 9 cm apart, what is the focal length of the convex mirror?

CHAPTER XLI

REFRACTION—PRISMS—LENSES

550. Refraction: Snell's laws. We have seen in Sec. 218 that **refraction is the bending of a beam of waves when it passes from one medium into another in which its speed is different from its speed in the first medium.**

Refraction takes place in accordance with the *laws of Snell* (Fig. 178):

1. When a beam of light passes from one medium into another in which its speed is less than in the first, it is bent toward the normal; and vice versa.

2. The ratio of the sine of the angle of incidence to the sine of the angle of refraction is a constant for the same two media.

3. The incident ray, the refracted ray, and the normal to the surface lie in a plane.

551. Index of Refraction. The **relative index of refraction** ${}_1\mu_2$ of a second medium with respect to a first medium has been defined as the ratio of the speed v_1 of the wave in the first medium to its speed v_2 in the second medium.

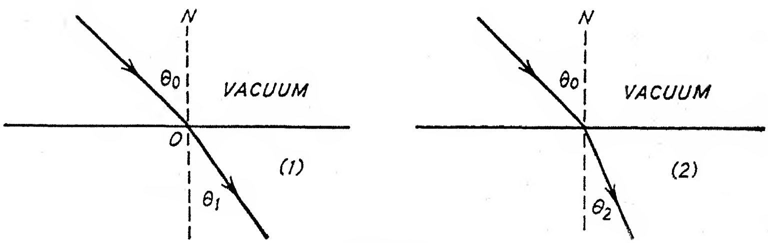


FIG. 530

In Art. 220, it was shown by Huyghens' principle that we should expect refraction and that

$$ {}_1\mu_2 \equiv \frac{v_1}{v_2} = \frac{\sin i}{\sin r} \qquad (417) $$

for all types of waves.

The **absolute index of refraction** μ_1 of a medium (1) is its index relative to a vacuum, i.e., when the beam of waves passes from a vacuum into the medium in question, Fig. 530.

Let v_0 be the speed of the waves in a vacuum, v_1 their speed in a medium (1) and v_2 their speed in medium (2). Then by Eq. (417), the absolute index of refraction for medium (1) is

$$\mu_1 \equiv \frac{v_0}{v_1} = \frac{\sin \theta_0}{\sin \theta_1} \tag{a}$$

and for medium (2) is

$$\mu_2 \equiv \frac{v_0}{v_2} = \frac{\sin \theta_0}{\sin \theta_2}. \tag{b}$$

Dividing (b) by (a),

$$_1\mu_2 \equiv \frac{v_1}{v_2} = \frac{\mu_2}{\mu_1} = \frac{\sin \theta_1}{\sin \theta_2} \tag{417a}$$

from which

$$\mu_1 \sin \theta_1 \equiv \mu_2 \sin \theta_2 = \mu_3 \sin \theta_3 = \text{const.} \tag{418}$$

This relation between absolute indices of refraction and the corresponding angles of refraction is equivalent to Snell's laws and is preferred by many physicists.

Unless otherwise specified, **indices of refraction relative to air are used.** The absolute index for air at standard conditions is 1.0002918 for light having the wave length of the *D*-line of sodium (λ = 5893 Å).* From this, other absolute indices may be calculated by means of Eq. (417a).

An important relation may readily be deduced from the definition of Eq. (417). Suppose a beam of light to pass successively through the media (1), (2), and (3) of Fig. 530a, and back into (1). (The surfaces of separation are not necessarily parallel for this proposition.)

Then, by the definition of index of refraction,

$$_1\mu_2 = \frac{v_1}{v_2}$$

$$_2\mu_3 = \frac{v_2}{v_3}$$

$$_3\mu_1 = \frac{v_3}{v_1}$$

* The Symbol Å means 1 Ångstrom unit of length = 10^{-8} cm.

from which, on multiplying the right and left members, respectively,

$$({}_1\mu_2)({}_2\mu_3)({}_3\mu_1) = \frac{v_1}{v_2} \times \frac{v_2}{v_3} \times \frac{v_3}{v_1} = 1. \tag{419}$$

As an application of this relation, consider a beam of light that passes from air (1) into glass (2) and back into air.

Then, by Eq. (419),

$${}_1\mu_2 \times {}_2\mu_1 = 1.$$

Therefore,

$${}_2\mu_1 = \frac{1}{{}_1\mu_2}. \tag{420}$$

That is, the index of refraction of air relative to glass is the reciprocal of the index of glass relative to air.

By again applying Eq. (419), we may deduce the index of refraction of one substance (say, water) with respect to another (glass) from data already known, without making the actual experimental determination. Thus,

Index of refraction of flint glass * (2) relative to air (1) $= {}_1\mu_2 = 1.65$

Index of refraction of water (3) relative to air (1) $= {}_1\mu_3 = 1.33$

Hence,

Index of refraction of air (1) relative to water (3) $= {}_3\mu_1 = \dfrac{1}{1.33} = .75$

and by Eq. (419), $({}_1\mu_2)({}_2\mu_3)({}_3\mu_1) = 1.65({}_2\mu_3)(0.75) = 1$

Index of refraction of water relative to flint glass $= {}_2\mu_3 = \dfrac{1}{1.65 \times 0.75} = 0.81.$

552. Refraction through plane parallel plates. When a ray of light passes through one or more layers of different media having plane parallel surfaces of separation and emerges again into the first medium, the emergent ray is parallel to the incident ray (Fig. 530a). This is easily shown as follows:

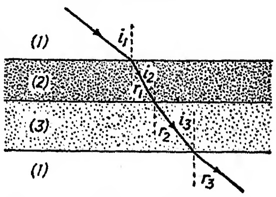


FIG. 530a. Refraction through Parallel Plates

By Eq. (418),

$${}_1\mu_2 = \frac{\sin i_1}{\sin r_1}$$

$${}_2\mu_3 = \frac{\sin i_2}{\sin r_2}$$

* Contains lead.

$$_3\mu_1 = \frac{\sin i_3}{\sin r_3}$$

and by Eq. (419),

$$({}_1\mu_2)({}_2\mu_3)({}_3\mu_1) = \frac{\sin i_1}{\sin r_1} \times \frac{\sin i_2}{\sin r_2} \times \frac{\sin i_3}{\sin r_3} = 1.$$

But since the planes are parallel, the normals are parallel also and therefore

$$r_1 = i_2 \quad \text{and} \quad r_2 = i_3 \quad (\text{alt. int. } \measuredangle \text{ of } \| \text{ lines}).$$

Hence,

$$\frac{\sin i_1}{\sin i_2} \times \frac{\sin i_2}{\sin i_3} \times \frac{\sin i_3}{\sin r_3} = 1$$

$$\sin i_1 = \sin r_3$$

$$\angle i_1 = \angle r_3.$$

Therefore the emergent ray is parallel to the incident ray.

553. Index of refraction dependent upon wave length. In Sec. 538 we saw that, except in a vacuum, the speed of light depends upon its wave length (color). Hence the index of refraction also depends upon the wave length.

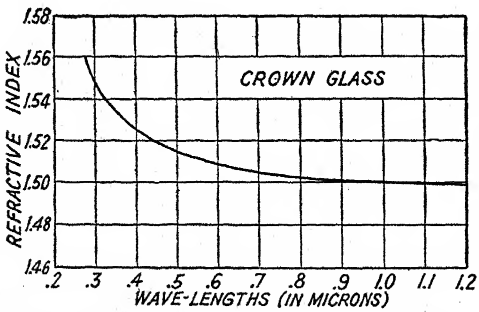


FIG. 531. Dispersion Curve for Crown Glass

In precise work, therefore, it is necessary to state the wave length of the light used in making the determination. For example, the index of refraction of rock salt is 1.54745 for the green line of mercury ($\lambda = 5461$ Å); but it is 1.54432 for the yellow line of sodium ($\lambda = 5893$ Å).

The index is usually greater for short wave lengths (blue end of the spectrum) than for long wave lengths (red end of the spectrum). This is shown in Fig. 531, which is the "dispersion curve" for a certain specimen of crown glass.*

* Does not contain lead.

Since temperature and pressure affect the properties of most materials, they also affect the index of refraction, and must therefore be stated, particularly in the case of liquids and gases.

554. Refraction by prisms: Dispersion. A prism is a body bounded by three or more planes that intersect in parallel lines. In the study of light, we are usually concerned with but one angle of the prism. This is known (Fig. 532) as the **refracting angle** ***A*****,** and its edge as the **refracting edge.**

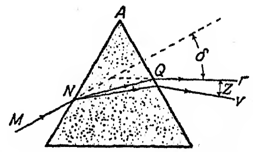


FIG. 532. Dispersion by a Prism

If the prism is denser than the surrounding medium, a ray of light is bent toward the normal on entering the prism, and away from the normal on leaving it, in accordance with Snell's laws. The angle δ between the entering ray MN and the emerging ray Qr is called the **angle of deviation** for the red ray.

Since light of short wave lengths (violet) travels more slowly in glass than does light of long wave lengths (red), the index of refraction for the former is greater than for the latter. Hence the violet light v is deviated through a greater angle than is the red ray r.

This results in the spreading out of the incident light into its constituent colors (wave lengths). The phenomenon is called **dispersion;** and the angle z between the extreme red and violet rays is the **angle of dispersion.**

The band of colors obtained is called a **spectrum.** A spectrum is frequently seen when sunlight passes through one of the angles of a cut-glass dish, or the beveled edge of a mirror.

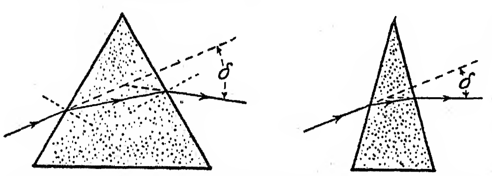


FIG. 533. Deviation Depends upon Refracting Angle

The angle of deviation will be greater, the greater the index of refraction. It will depend also upon the size of the refracting angle A. Within limits, the greater the refracting angle, the greater will be the deviation, as shown in Fig. 533.

If a prism is placed on the turntable of a spectrometer (Sec. 607) and rotated, it is found that there is a certain **unique** position of the prism for which the **angle of deviation is a minimum**

for a given color. Examination of the divided circle will then show that for this position, the entering ray and the emerging ray make equal angles α and β with the corresponding faces of the prism (Fig. 534). This fact may be readily demonstrated theoretically.* A prism is usually set for minimum deviation because its resolving power is then a maximum.

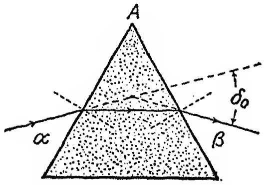


FIG. 534. Ray Traversing Prism with Minimum Deviation

One of the best methods of determining the index of refraction of a prism for a given wave length is by using the value of its angle of minimum deviation δ_0. If μ is the index of refraction and A the refracting angle, the relation * is:

$$\mu = \frac{\sin \frac{1}{2}(A + \delta_0)}{\sin \frac{1}{2}A}. \tag{421}$$

555. Lenses. A lens is a piece of material one or more of whose surfaces is curved, and whose density is different from that of the surrounding medium. The most common forms are bounded by surfaces of revolution and are classified as follows (Fig. 535):

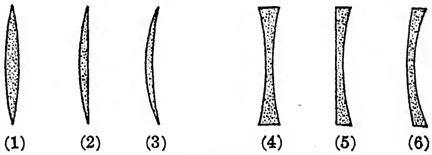


FIG. 535

(*a*) *Converging lenses*, which are thicker at the center than at the circumference and have the names: (1) Double convex, (2) Plano-convex, (3) Concavo-convex.

(*b*) *Diverging lenses*, which are thinner at the center than at the circumference and have the names: (4) Double concave, (5) Plano-concave, (6) Convexo-concave.

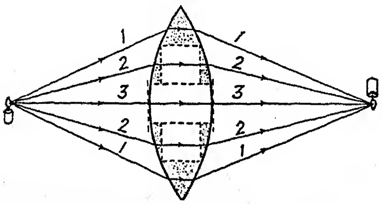


FIG. 536. A Lens as a Series of Prisms

The effect of a lens may be understood by thinking of it as a series of prisms (Fig. 536). The topmost prism has the greatest refracting angle, and hence ray 1 is deviated most. The next lower

* Edser, *Light for Advanced Students* (New York, The Macmillan Company, 1925), p. 57.

section consists of a parallel plate, which does not bend the ray, between two prisms (dotted) of small angle. Hence ray 2 is deviated less than ray 1. At the center, ray 3 is passing through a parallel plate and is not deviated. In the lower half of the lens, the same deviations take place in reverse order. Hence this lens causes the rays to converge.

From the standpoint of waves, the action of converging and diverging lenses is illustrated in Fig. 537. The former is thicker at its center than at its edge. Therefore the portion of the wave front passing through the center is retarded more than other portions, the least retardation being at the edge. Hence the emerging wave *converges* to a focus at F.

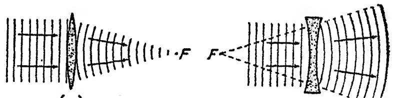


FIG. 537. Converging and Diverging Lenses

In Fig. 537b, just the reverse takes place. The part of the wave front passing through the center of the lens is retarded least, the retardation increasing gradually out to the periphery. Hence the emerging waves are *diverging;* i.e., the focus is behind the lens (virtual) at F.

If the material of the lens were less dense than the surrounding medium (e.g., an air cavity in a cube of glass), wave advancement would replace retardation, and the terms "converging" and "diverging" would be reversed.

556. Definitions. The following definitions apply to lenses, and refer to Fig. 538.

The **principal axis** Bb of a lens is the straight line passing through the centers of curvature of its two surfaces.

The **principal focus** is that point on the principal axis through which rays of light that before refraction were parallel to the principal axis pass after refraction. F_1 is the principal focus for parallel light from the left; and F_2, for parallel light from the right.

The **optical center** C is that point on the principal axis through which a ray of light may pass without suffering a permanent change in direction. A ray passing through the optical center is given a slight lateral displacement, since in effect it is traversing a parallel plate; but this displacement is generally negligible in the case of thin lenses.

The optical center for thin lenses, both of whose surfaces have

the same radius of curvature, is at the geometrical center of the lens. For the plano-lenses, it is at the point where the principal axis meets the curved face. For the other types it is at a point on the principal axis outside the lens on the concave side.

In the case of thick lenses, *no single point* plays the role of optical center: it is necessary to use two points called *nodal points*.

The **focal length** f is the distance CF_1 from the optical center to the principal focus. The equivalent focal length of a combination of lenses is the focal length of a single lens that would (theoretically) replace the combination.

The **aperture** of a lens is the diameter of the uncovered part.

Conjugate foci are any two points so situated that light emanating from one and passing through the lens is brought to a focus at the other, and vice versa. A and a are conjugate foci; either may be considered to be the image of the other.

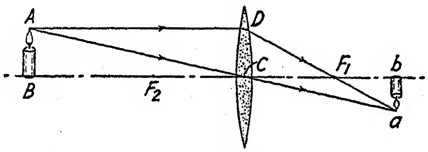


FIG. 538. Construction of Image Formed by Lens

557. Construction of image formed by thin lens.* From any point A of the object (Fig. 538), we draw two rays: one parallel to the principal axis, and the other through the optical center. After refraction, the former goes through the principal focus F_1; the latter continues unchanged in direction. The point a where these two refracted rays meet is the image of the point A.

If the rays actually meet, the image is *real* and can be caught on a screen. If they do not actually meet, but appear to meet when prolonged backward, the image is *virtual* and cannot be caught on a screen.

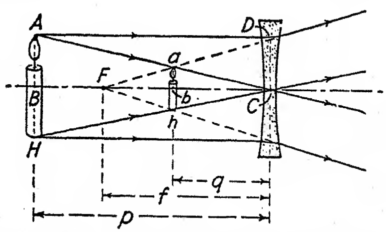


FIG. 539. Image Formed by Diverging Lens

With a single lens and a luminous object such as a candle, a **real image** is always *inverted;* and a **virtual image** is always *erect.*

All rays emanating from A and passing through the lens converge to a, except certain marginal rays that will be discussed later.

* For a discussion of thick lenses, see Edser, *Light for Students*, p. 135.

For a **diverging lens,** the construction is exactly the same as the foregoing; but the image is always virtual (Fig. 539).

558. Location of image: General lens formula.

CASE I: When the object is at a distance from a *converging* lens greater than the focal length (Fig. 540).

We know that $\triangle ABC$ is similar to $\triangle abC$ (rt. $\triangle$s having acute $\angle$s equal).

$$\frac{AB}{ab} = \frac{p}{q} \quad \text{(homologous parts of similar figures).} \quad \text{(a)}$$

Also, $\triangle DCF$ is similar to $\triangle abF$ (same reason as above).

$$\frac{DC}{ab} = \frac{CF}{bF} = \frac{f}{q-f}. \quad \text{(b)}$$

But,

$$DC = AB \quad \text{(opposite sides of rectangle)}$$

$$\frac{AB}{ab} = \frac{f}{q-f}. \quad \text{(c)}$$

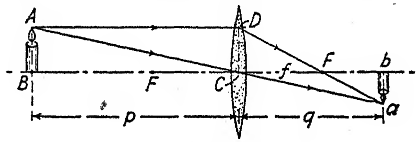


FIG. 540

From Eqs. (a) and (c),

$$\frac{p}{q} = \frac{f}{q-f}. \quad \text{(d)}$$

Clearing,

$$pq - pf = qf$$

or,

$$qf + pf = pq.$$

Dividing through by pqf,

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}. \quad \text{(e)}$$

CASE II: When the object is at a distance from a *converging* lens less than the focal length (Fig. 541).

The derivation is the same as in Case I as far as Eq. (b), which here becomes:

$$\frac{DC}{ab} = \frac{f}{q+f}. \quad \text{(f)}$$

Therefore,

$$\frac{p}{q} = \frac{f}{q+f}$$

$$pq + pf = qf.$$

Dividing by pqf,

$$\frac{1}{p} - \frac{1}{q} = \frac{1}{f}. \quad \text{(g)}$$

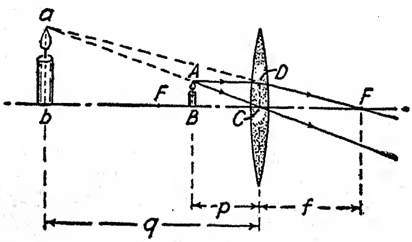


FIG. 541

CASE III: When the lens is *diverging* (Fig. 539).

Again the derivation is the same as Case I down to Eq. (b), which in this case is:

$$\frac{DC}{ab} = \frac{f}{f-q} \tag{h}$$

$$\frac{p}{q} = \frac{f}{f-q}$$

$$pf - pq = qf.$$

Dividing by pqf,

$$\frac{1}{p} - \frac{1}{q} = -\frac{1}{f}. \tag{i}$$

In deriving the three formulas above, we took all distances to be positive, just as they were in the various figures. Inspection of Eqs. (e), (g), and (i) will show that if we adopt the following **convention of signs,** the last two forms become the same as the first one.

1. f is to be taken + for converging lenses, and − for diverging lenses.

2. p is to be taken + when the object is real, and − when the object is virtual. (See Sec. 560.)

3. q is to be taken + when the image is real, and − when the image is virtual.

When these rules are observed, Eqs. (g) and (i) take the same form as (e), which is therefore called the

General lens formula:

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}. \tag{422}$$

It will be noticed that this is exactly the same as the general mirror formula, Eq. (415), and that the convention of signs is likewise the same as for mirrors.

559. The lens maker's equation. When a lens maker sets out to make a lens, he must know the proper radii R_1 and R_2 to which to grind the surfaces so that the lens will have the required focal length. The focal length will obviously depend also upon the index of refraction μ of the glass. It may be shown* by the

* Preston, *Theory of Light* (New York, The Macmillan Company, 1928), p. 86.

method of Sec. 545 that

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q} = (\mu - 1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{423}$$

which is therefore called the "lens maker's equation."

The quantities are all taken positive for the typical double convex lens when object and image are both real. For other lenses they are taken positive when on the same side of the lens, as in the case of the typical double convex lens; otherwise negative.

560. Lens combinations. Virtual object. When two or more thin lenses are used in combination, the position and size of the

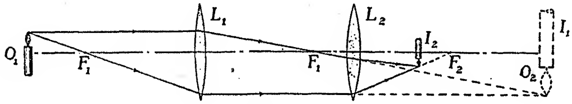


FIG. 542

image formed by the first lens are found as in Sec. 557. This image then becomes the object for the second lens, and the process is repeated. If the rays from the first lens enter the second lens before they meet in the first image, that image becomes the **virtual object** for the second lens, as is shown in Fig. 542.

Here the object O_1 is at a distance ($p_1 = 18$ cm) from a converging lens L_1 of focal length ($f_1 = 12$ cm). By Eq. (422), it is found that the image I_1 would be at the distance ($q_1 = 36$ cm) from L_1. Before this image can be formed, however, the rays enter the lens L_2, which has the focal length ($f_2 = 10$ cm) and is at the distance ($d = 16$ cm) from L_1.

Hence the image I_1 is not actually formed and could not be caught on a screen. So when it is used as the object O_2 for the second lens, it is called the "virtual object." But the general lens formula holds true for this case provided the **distance p_2 from the virtual object to the second lens is taken negative.**

In this case, the virtual object distance is:

$$p_2 = -(q_1 - d) = -(36 - 16) = -20 \text{ cm}.$$

Using this in Eq. (422),

$$\frac{1}{10} = \frac{1}{-20} + \frac{1}{q_2}$$

$$q_2 = 6.67 \text{ cm}.$$

561. Relative size of image and object. From Eq. (a) of Sec. 558,

$$\frac{ab}{AB} = \frac{q}{p}.$$

But ab and AB are the size of image and object, respectively; and q and p are their respective distances from the optical center of the lens. Hence, in words,

$$\text{Magnification} \equiv \frac{\text{Size of image}}{\text{Size of object}}$$
$$= \frac{\text{Distance of image from lens}}{\text{Distance of object from lens}}. \qquad (424)$$

It should be noted that this equation does *not* give the magnifying power of a lens, for the size of the image will be different for every different object distance. To define magnifying power, one of the distances must be fixed. This will be taken up in Chap. XLVI.

562. Real and virtual images by lenses. By applying the general lens formula, Eq. (422) and the construction of Sec. 557, the results will be found to fall into six cases.

Single Converging Lens with Real Object:

1. When the object distance is greater than $2f$, the image is real, inverted, and smaller than the object.
2. When the object distance is $2f$, the image is real, inverted, and of the same size as the object.
3. When the object distance is less than $2f$, but greater than f, the image is real, inverted, and larger than the object.
4. When the object distance is f, i.e., when the object is at the principal focus, the refracted rays are parallel and hence no image is formed in finite space.
5. When the object distance is less than f, the image is virtual, erect, and larger than the object.

Single Diverging Lens with Real Object:

6. For all object distances, the image is virtual, erect, and smaller than the object.

563. Spherical aberration. If the surfaces of a lens are spherical, rays such as 1 and 4 (Fig. 543), which strike the lens at a consid-

erable distance from the center, are brought to a focus J closer to the lens than are rays 2 and 3, which strike the lens near its center and are brought to a focus at F.

This defect of spherical lenses is known as *spherical aberration.* It may be partly overcome by "stopping down" the lens, i.e., by screening the outer part so that only the central portion is used. This method reduces the aperture and the "speed" of the lens. (See Sec. 617.)

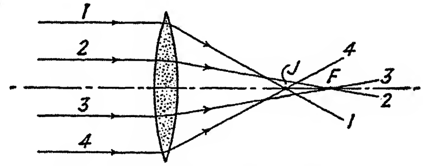


Fig. 543. Spherical Aberration

By properly decreasing the curvature of a lens as its diameter increases, spherical aberration may be corrected for a single point (only) on the axis, so that rays of one wave length from points in a plane normal to the principal axis at that point will be brought to a focus in a plane. Such lenses are said to be *aplanatic.* But the correction is usually made in connection with that for chromatic aberration and astigmatism by a combination of lenses made of different glasses (Fig. 546).

564. Chromatic aberration. It will be seen (Sec. 596) that on passing through a prism, a beam of light suffers dispersion into its constituent colors. A lens, essentially a series of prisms, also produces dispersion. The short waves (violet) are brought to a focus F_v closer to the lens than the focus F_r of the longer (red) rays (Fig. 544).

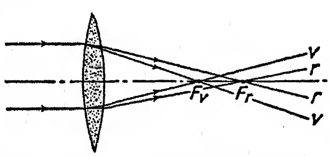


Fig. 544. Chromatic Aberration

This defect of a lens is known as *chromatic aberration.* It may be corrected for two chosen colors by building up a doublet of two different kinds of glass. Thus, a converging lens of crown glass (Fig. 545a) is designed, producing a certain dispersion rv and a certain deviation δ_C. A diverging lens (Fig. 545b) of flint glass is also made, producing the same amount of dispersion vr and a deviation δ_F, both opposite to those of the crown lens.

When these two component lenses are placed together, as in Fig. 545c, the dispersion of the second neutralizes that of the first; but since δ_C is greater than δ_F, there remains an outstanding de-

viation δ. Such a lens, corrected for two colors, is called an **achromatic** lens.

By using three or more glasses, correction may be made for chromatic aberration of three different wave lengths, which prac-

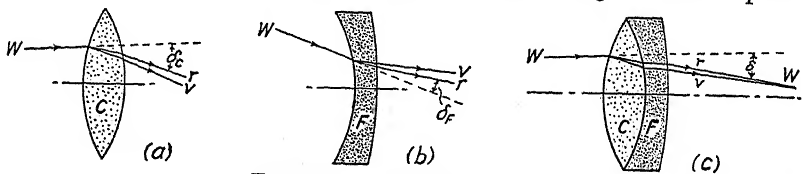


Fig. 545. Achromatic Doublet

tically eliminates the defect. Lenses corrected for three different wave lengths are called **apochromatic lenses.**

565. Astigmatism of Lenses. Astigmatism is the phenomenon in which light from a point source *P* (Fig. 546a) is brought to a

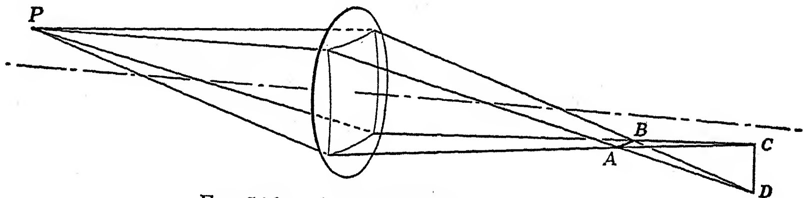


Fig. 546a. Astigmatic Images of a Point

focus along two lines *AB* and *CD* some distance apart instead of being focused at a single point.

It occurs with the lens of some eyes because the curvature of the surface is different in different axial planes, and it may be corrected by a combination of cylindrical and spherical lenses so as not to be noticeable.

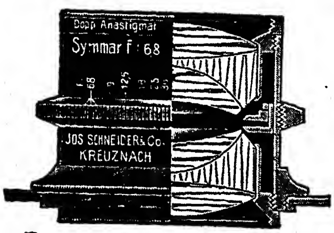


Fig. 546b. Anastigmatic Lens

But even if the lens surface is perfectly spherical, astigmatism occurs when the point source is not on the principal axis of the lens. In this case, it may be corrected by a combination of lenses made from different kinds of optical glass.

A lens corrected for astigmatism as well as for spherical and chromatic aberration is called an **anastigmatic lens,** or an **anastigmat** (Fig. 546b).

566. Lens power. In Fig. 537, we saw that a plane wave (curvature zero) was brought to a focus at F. That is, the curvature of the wave as it emerges from the lens is $1/f$, where f is the focal length.

Hence $1/f$ represents the curvature impressed upon the wave by the lens, or the ability of the lens to produce curvature. This ability of a lens to produce curvature is called by opticians the *lens power*.*

The unit of lens power is called the **diopter. One diopter is the power of a lens whose focal length is 1 meter.** Hence,

$$\text{Lens power (in diopters)} = \frac{1}{f}\ \text{(in meters)}. \tag{425}$$

For thin lenses in contact, lens powers are additive. That is, if f is the focal length of the combination,

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} + \frac{1}{f_3} + \cdots. \tag{426}$$

567. Total reflection. When a ray of light AO (Fig. 547) strikes the surface of separation XY between two media in which its velocity is different, part of the light OA'' will in general pass into the second medium, while the remainder OA' will be reflected in the first medium.

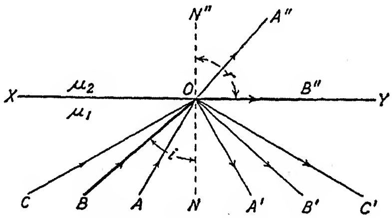


Fig. 547. Total Reflection

If the index of refraction μ_2 of the second medium is less than the index μ_1 of the first medium (both with respect to air), then, as the angle of incidence is increased, the angle of refraction will increase also until a position BO of the incident ray will finally be reached for which the angle of refraction is 90°. The refracted ray OB'' will then lie along the surface of separation.

If the angle of incidence is further increased, no part of the light will emerge into the second medium: **it will be totally reflected within the first medium,** as shown by rays CO and OC'.

This phenomenon is called **total reflection.** The angle of inci-

* This term should not be confused with magnifying power.

dence for which the angle of refraction is 90° is called the **critical angle** of the first medium with respect to the second.

In brief, when a beam of light in a first medium strikes the surface of a second medium whose index of refraction is less than that of the first medium, the beam will be totally reflected in the first medium if the angle of incidence is greater than the critical angle.

The numerical value of the critical angle for any medium 1 with respect to any other medium 2 is easily deduced from Snell's laws, as follows.

Since for the critical angle of incidence, the angle of refraction is 90°, we have, by Sec. 551,

$$ {}_1\mu_2 = \frac{\sin i}{\sin 90^\circ} = \sin i \tag{427}$$

where ${}_1\mu_2$ is the index of refraction of the second medium relative to the first, according to the definition in Sec. 221, and sin i is the **sine of the critical angle** of the first medium relative to the second.

Consider, for example, the case of light passing from water into air. Here water is the first medium and air is the second, so that ${}_1\mu_2$ is the index of refraction of air with respect to water. The index of refraction of water with respect to air is 1.33, which in this case would be ${}_2\mu_1$. Hence, by Eq. (420),

$$ {}_1\mu_2 = \frac{1}{{}_2\mu_1} = \frac{1}{1.33} = 0.75. $$

Therefore

$$ \sin i = 0.75 $$

and

$$ i = 48^\circ\ 35'. $$

Hence the critical angle of water with respect to air is 48° 35′.

Some other critical angles with respect to air are:

Crown Glass......	41°
Flint Glass.......	37°
Diamond.........	24° 25′

The critical angle for diamond is so small that total reflection is easily secured over a wide range. This accounts for the exceptional "sparkling" brilliance of the diamond.

It will be observed also that the critical angles for both crown and flint glass are less than 45°. Hence a 45° prism of either of these glasses will serve to turn a ray of light through a right angle by total reflection (Fig. 548a), or to invert an image (Fig. 548b).

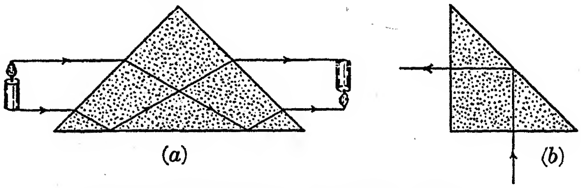


FIG. 548. Applications of Total Reflection

As a means of changing the direction of a beam of light, total reflection is not 100% efficient on account of the absorption of light in the first medium. Yet it is in general superior to the use of mirrors. Consequently, totally reflecting prisms are very extensively used in optical instruments, e.g., in the prism binocular (Fig. 549).

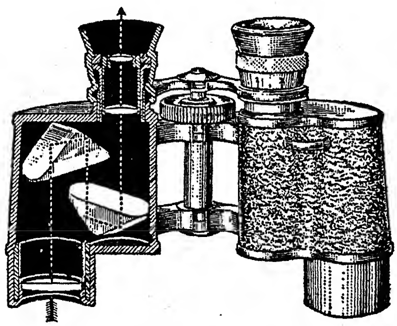


FIG. 549. The Prism Binocular. (From Millikan, Gale and Edwards, *First Course in College Physics*, courtesy of Ginn and Co.)

We are now able to explain the **mirage,** mentioned in Sec. 532. On a very hot day, the air is hottest near the surface of the ground and gets cooler gradually as we go upward. Hot air is optically less dense than cool air. Hence a ray of light inclined downward from an object is passing continually from optically denser to optically rarer media, until it finally undergoes total reflection and is bent upward toward the eye of the observer. This is shown by the arrows on the heavy line of Fig. 513.

Since the eye "sees" an object in the direction from which the light enters the eye, there will appear an inverted image of the tree. The observer may therefore infer the presence of water, in which reflections more commonly occur. The same phenomenon occurs at sea, with the conditions reversed: mariners sometimes see in the sky the inverted image of a vessel that is actually below the horizon.

PROBLEMS

1. If a beam of light makes an angle of 40° with the surface of the water in a tank, what is the angle of refraction? The angle of deviation?

2. The index of refraction of crown glass with respect to air is 1.5, and that of water with respect to air is 1.33. Compute the index of water with respect to crown glass, and of crown glass with respect to water.

3. The index of refraction of benzene with respect to air is 1.501. What is the index of benzene with respect to water?

4. What is the angle of minimum deviation for a 60° prism of flint glass ($\mu = 1.65$)?

5. If the angle of minimum deviation of a certain 60° prism for sodium light is 38° 56′, what is the index of refraction of the glass?

6. A converging lens of 20-cm focal length is placed 15 cm from an object. Where will the image be formed? If the object is 4 cm tall, what is the size of the image? Illustrate by diagrams.

7. What should be the focal length of a lens placed 15 cm from a plate in order to photograph an object 6 m distant?

8. An object 3 cm high stands on the optic axis and 10 cm from the optical center of a plano-concave lens ($f = -6$ cm). Locate the image by means of a scale drawing and check by the lens formula.

9. What should be the focal length of a lens in a camera in order to take a 12-cm picture of a 15-m building at a distance of 40 m from the camera?

10. What should be the focal length of a lens in a camera in order to take an 8-cm picture of a 10-m building at a distance of 30 m from the camera?

11. A projecting lantern is to produce a magnification of 40 diameters at a distance of 50 ft. Find the distance of the lens from the slide, and the focal length of the lens.

12. A projection screen is 8 ft wide and a lantern slide 3 in. wide. If the screen is 32 ft from the lantern, what must be the focal length of the projection lens if the image is to cover the screen exactly?

13. A double concave lens of 40-cm focal length is to be ground from crown glass ($\mu = 1.5$). If both surfaces are to have the same curvature, what must the radius be?

14. A biconvex lens, both of whose faces have the same curvature, is to be made from flint glass ($\mu = 1.65$). Compute the proper radii for a focal length of 20 cm.

15. The radii of curvature of the two faces of a double convex lens are 40 cm and 60 cm, respectively. If flint glass ($\mu = 1.66$) is used, compute the focal length.

16. Two thin lenses have powers of 2 and 5 diopters, respectively. What is the focal length of each and of the combination when placed in contact? What is then the power of the combination?

17. A double convex lens has radii of curvature of 24 cm and 30 cm, respectively, and its index of refraction is 1.5. What is its power in diopters?

18. A negative lens has a power of 2 diopters. If an object is 40 cm from the lens, where is the virtual image?

19. Two positive lenses are 40 cm apart. The focal length of the first is 20 cm and of the second, 6 cm. An object is 60 cm in front of the first lens. Where will the final image from the second lens be?

20. Two positive lenses of focal lengths 30 cm and 10 cm, respectively, are 5 cm apart. If an object is 90 cm in front of the first lens, where will be the final image formed by the combination?

21. In order to find the focal length of a diverging lens, a converging lens ($f = 10$ cm) is placed 11 cm behind the diverging lens. The real image of an object 12 cm in front of the diverging lens is then 30 cm behind the converging lens. Find the focal length of the diverging lens.

22. If the index of refraction of a certain kind of glass is 1.55, what is the critical angle?

23. If the critical angle of ethyl alcohol relative to air is 48° 20′, what is its index of refraction?

24. The index of refraction of carbon bisulphide relative to air at 15°C is 1.63. What is the critical angle?

25. Find the velocity of light in water if the critical angle for water and air is 48°30′.

26. The index of refraction of diamond is 2.417 relative to air at 15°C. Find its critical angle. How does this enhance its value as a gem?

CHAPTER XLII

DIFFRACTION—INTERFERENCE

568. Diffraction has already been defined in Sec. 222 as the bending of waves into the geometrical shadow of an obstacle. Diffraction of light was discovered by Grimaldi of Bologna about 1616. It supplied precisely the evidence that Newton himself had declared necessary to substantiate the wave theory; but that great philosopher missed its significance, and undertook without success to explain the "diffraction bands" of Grimaldi by his emission theory (Fig. 550).

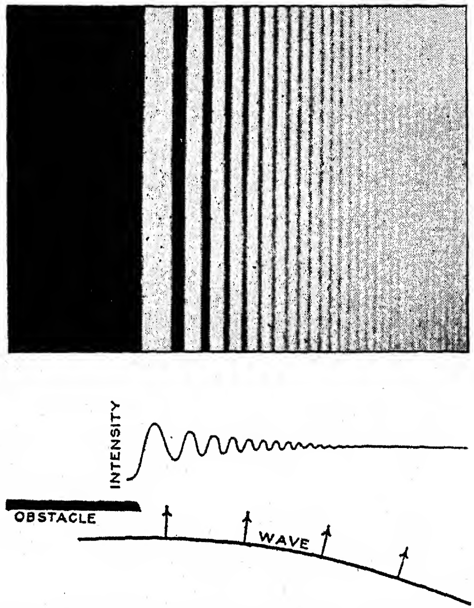


Fig. 550. Diffraction at Sharp Edge. (Courtesy Professor M. E. Huffard)

569. Young's experiment. The phenomenon of diffraction was not understood until 1801 when Thomas Young, professor of natural philosophy at the Royal Institution, London, demonstrated the interference of light by the following epoch-making experiment.

Light from a strong source such as a carbon arc L (Fig. 551) passes through a slit S in a screen AB and falls upon two slits S_1 and S_2 in a second screen CD, the three slits being accurately parallel.

In accordance with Huyghens' principle, slits S_1 and S_2 become new sources, from each of which a pencil of waves goes forward

as shown. Each pencil spreads out by diffraction into a fan-shaped region of disturbance. These two wave systems overlap, and where they meet crest on crest (in phase) there will be reinforcement; i.e., the amplitude will be twice as great and therefore the intensity four times as great as for the waves of either beam alone.

At other places, the two systems of waves meet crest on trough

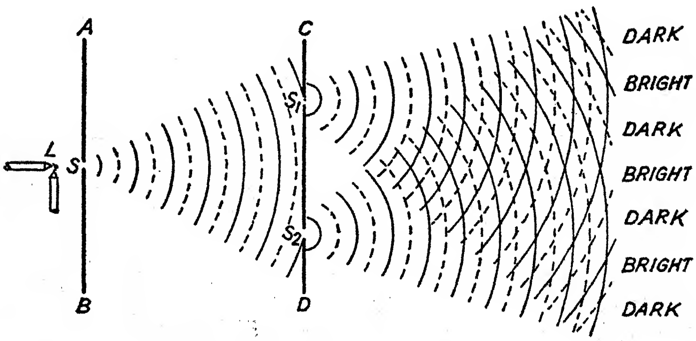


Fig. 551. Young's Experiment

(out of phase), and at these places the two waves neutralize each other; i.e., there is destructive interference with resulting zero intensity, or darkness.

It should be noted that interference can be obtained only when both beams originate from the same point source S.

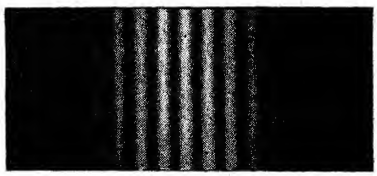


Fig. 552. Bands, or Fringes, in Young's Experiment. (Courtesy Professor H. E. White)

Since the sources are slits perpendicular to the plane of the page, the bright and dark regions will have the form of lines, or bands, parallel to the slits. They may be seen by looking from right to left along the center line of the arrangement, and are best observed by means of a low-power telescope. The appearance of the bands is shown in Fig. 552.

If a source of white light is used, the bands, or "fringes," on either side of the center are edged with color because of the difference in wave length, the longer (red) waves yielding bands that are more widely separated than those produced by the shorter waves. This is nicely shown by interposing screens of different

colors in the line of sight. The fringes formed by blue light are noticeably narrower than those formed by red light.

Only waves exhibit both reinforcement and destructive interference at the same time. Hence **the importance of Young's experiment is that it proves the wave nature of light.** But this was not recognized until about twenty years later, after the corroborative work of Arago and Fresnel. In the meantime, Young was the victim of popular disapproval on account of the vicious attacks of one Lord Brougham, a fanatical writer in the *Edinburgh Review*, who then held the public ear.

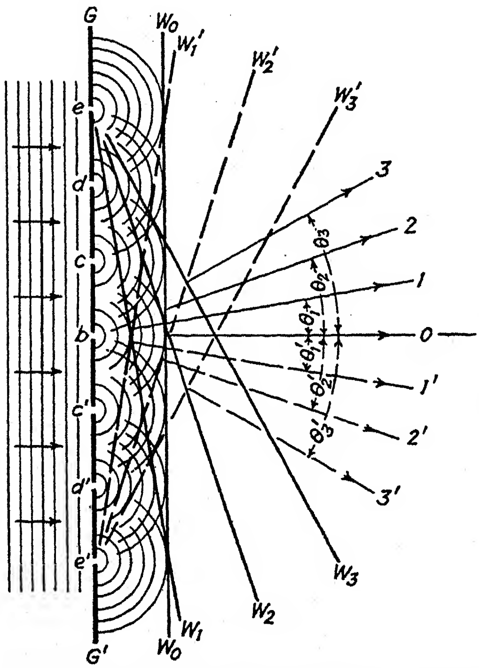


FIG. 553. Huyghens' Principle Applied to Grating

570. The diffraction grating. This very important device consists of a great number of narrow, parallel strips through which light may pass, or from which it may be reflected. The interference of the resulting diffracted beams of light breaks it up into its constituent colors, forming several spectra which are said to be of different orders.

The first gratings were made of fine wire by Joseph von Fraunhofer. Modern gratings, however, which sometimes have as many as 30,000 lines per inch drawn by a diamond point on transparent glass or on a metallic mirror, were first successfully ruled by Henry A. Rowland of the Johns Hopkins University.

The lines drawn by the diamond are rough and do not transmit or reflect light regularly. They correspond to the opaque parts of the grating GG' in Fig. 553, in which the principle of the grating is explained.

A beam of parallel rays of monochromatic light (i.e., light hav-

ing but one wave length) impinges normally upon a screen GG' in which there are a number of parallel slits e, d, c, b, c', d', e' perpendicular to the plane of the paper.

In accordance with Huyghens' principle, each slit becomes a new source from which wavelets go out as from the original source; and a new wave front is an envelope of a group of these wavelets, all of which are in the same phase.

It will be seen then that we may expect the original wave system to be broken up into several wave systems as follows, only a single wave front being drawn for each system:

(a) The **central-image** system w_0w_0, parallel to the incident beam.

(b) The **first-order** systems ew_1 and $e'w_1'$, whose directions of propagation are the rays $b1$ and $b1'$, making the angles θ_1 and θ_1', respectively, with the axis bo. In these directions, the waves from e and e' are one wave length behind those from d and d'; two wave lengths behind those from c and c', and so on. Consequently, all these waves meet crest on crest (in phase) along the wave fronts ew_1 and $e'w_1'$, respectively, and produce reinforcement.

(c) The **second-order** systems ew_2 and $e'w_2'$, whose directions of propagation are $b2$ and $b2'$, respectively. In these directions, a wave from e or from e' is two full wave lengths behind a wave from d or d', respectively, and so on. These waves meet in phase along the wave fronts ew_2 and $e'w_2'$, respectively, and produce reinforcement.

(d) Similarly, **third-, fourth-, fifth-,** etc., order systems are found in certain definite directions.

How these directions of the different orders of spectra are related to the wave length of the incident light and to the **grating space** a (i.e., the distance between the centers of adjacent slits) is easily seen in the enlarged drawing of Fig. 554a.

Here it will be observed that a wave just starting at c is in phase with one from b which has advanced one whole wave length λ; with one from c' which has advanced 2λ; with one from d' which has advanced 3λ; etc. These waves are therefore in phase along the wave front cs, and give the parallel beam of the **first order,** whose direction of propagation is $b1$. In isotropic media, the direction of propagation is normal to the wave front.

Hence, $\triangle csb$ is a rt. $\triangle$ and therefore,

$$\sin \theta_1 = \frac{\lambda}{a}$$

$$a \sin \theta_1 = \lambda.$$

Similarly for the second-order spectrum (Fig. 554b),

$$\sin \theta_2 = \frac{2\lambda}{a}$$

$$a \sin \theta_2 = 2\lambda.$$

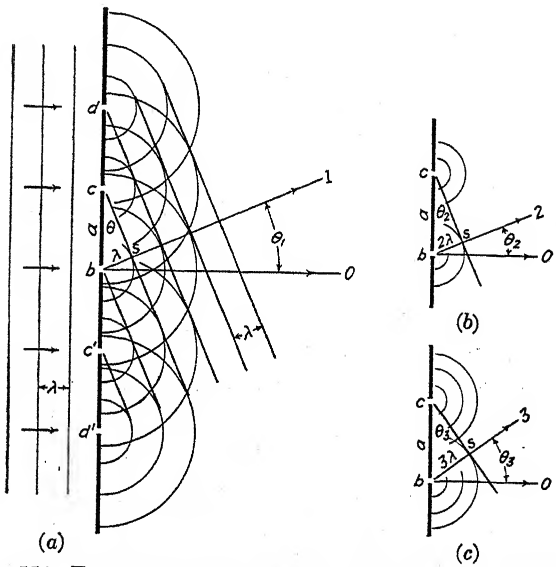


FIG. 554. Formation of First, Second, and Third Order Spectra

And for the third order (Fig. 554c),

$$\sin \theta_3 = \frac{3\lambda}{a}$$

$$a \sin \theta_3 = 3\lambda.$$

From these special cases, we can state the

General law of the diffraction grating:

$$a \sin \theta = n\lambda \tag{428}$$

where a is the grating space;
θ is the angle of deviation from the central beam;
n is the order of the spectrum; and
λ is the wave length of the light.

From this it is seen that for each value of n, i.e., in each order, the sine of the angle of deviation is proportional to the wave length. Consequently, in each order we get a band of colors, or spectrum, in which are represented all the constituent colors (wave lengths) of the incident light. When the incident light is normal to the grating, the orders are in general symmetrically arranged on opposite sides of the axis, or of the direction of the undeviated ray; but this is not always the case.

For small changes of θ, the change in angle is nearly proportional to the change in wave length. In any case there is a definite geometric relation between wave lengths and their position in a spectrum produced by a grating. Hence a spectrum produced by a grating is called a **normal spectrum** (see Sec. 613).

By inspection of Eq. (428), it will be seen that all the quantities except λ can be easily measured.

a is known from the design of the ruling engine on which the grating is made.
θ is directly observed on the scale of the spectrometer.
n is the order of the spectrum in which the observation is made.

Hence the equation may be solved for λ.

The grating, therefore, gives an independent and highly accurate method for the absolute determination of the wave length of light. This is not true of the prism.

571. Colors of thin films: Newton's rings. In the two preceding sections, interference took place between diffracted beams. But light that has not undergone diffraction exhibits interference equally well. It is only necessary that the interfering waves originate at the same source.

The colors seen on soap bubbles, films of oil on water, mother of pearl, thin layers of oxide on metal, etc., are due to the interference of two beams of light reflected at the two surfaces of the films in question.

To study the phenomenon, Newton (*Optics* II, 1704) placed a convex lens of slight curvature on a plane glass surface and ob-

served the series of colored rings produced when the device is illuminated with white light. He tried without success to explain them by his corpuscular theory of light. The explanation is easy on the basis of Huyghens' wave theory, using a simple form of interferometer devised by Professor Gale of the University of Chicago.*

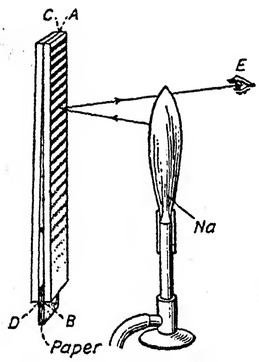


FIG. 555. Simple Interferometer. (Courtesy Ginn and Co.)

Two strips of plate glass (Fig. 555) are separated at the bottom by a piece of thin paper and held together at the top by a clamp (not shown). Between them there is a thin film, or wedge, of air whose thickness at the top is zero and at the bottom is the same as that of the paper.

If a source of monochromatic light, such as a gas burner impregnated with sodium chloride, is placed in front of the glass plates, the eye at E will observe narrow bright and dark bands across the plate, as shown.

The light striking the plates is reflected partially at each of the four glass-air boundaries; but we need to consider now only the surfaces which enclose the wedge of air, i.e., AB and CD (Fig. 556). The individual atoms of the luminous sodium are the sources of light.

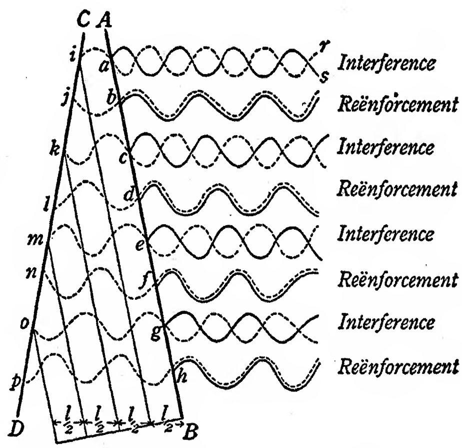


FIG. 556. Formation of Bands or Fringes. (Courtesy of Ginn and Co.)

Consider a wave radiated by *one* of these atoms. Part of its energy is reflected at a along the full wavy line as. The rest of it is reflected at i along the dotted wavy line ir. If the distance ai is half a wave length ($\lambda/2$), the part reflected at i travels a whole wave length farther than that reflected at a. We would therefore

* H. G. Gale, in the *American Machinist* (July 11, 1901).

expect them to meet in phase and reinforce. But the reflection at i takes place in a rarer medium (air) against a denser medium (glass). In such cases there is always a change of 180° in phase.* Hence, where the thickness of the air wedge is $\lambda/2$, the two reflected waves are 180° out of phase; and, since they originate at the same luminous atom, they annul each other, producing a dark band.

Similarly, at bj, where the thickness of the air wedge is $\frac{3}{4}\lambda$, the path difference is $\frac{3}{2}\lambda$, and the change of 180° at j makes the two reflected waves coming to the eye in phase, producing reinforcement of the original light.

Thus, as we go from top to bottom, we get a **dark band** wherever the thickness of the air wedge is an **even number** of quarter wave lengths ($\lambda/4$), and a **bright band** wherever the thickness is an **odd number** of quarter wave lengths. If the device is illuminated by white light, the bands are colored violet, green, yellow, etc., as the thickness of the air film increases.

In the case of Newton's apparatus, the bands are obviously rings, since the loci of equal air-film thickness are circles.†

PROBLEMS

1. A grating having 6000 lines/cm is set up on a spectrometer and the direction of the second-order image of the slit is found to make an angle of 45° with the normal. What is the wave length of the light used?

2. A grating is set up at 100 cm from a meter-stick on the perpendicular at its middle point. A lithium light is placed behind the slit at the center of the meter-stick. On looking through the grating, the first-order images of the slit for 6708 Å are seen 88 cm on each side of the slit. What is the grating space and the number of lines per cm on the grating?

3. A grating having 14,500 lines/in. is placed 1 m from a scale on the perpendicular to the scale at its middle point. If sodium light is used behind a slit at the center of the scale, where will the first-order diffraction image be seen?

4. A grating having 15,000 lines/in. is placed 80 cm from a scale on the perpendicular at its middle point. If sodium light is used, where will the diffracted image of the second order be seen on the scale?

5. A grating having 15,000 lines/in. is placed 80 cm from a scale on the perpendicular at its middle point. If sodium light is used behind a slit at the center of the scale, where will the third-order diffraction image be seen?

6. If a beam of sodium light (5893 Å) is divided, half being passed through a column of water 100 cm long and the other half through the air, how long will

* Edser, *Light for Students* (New York, The Macmillan Co., 1925), p. 283.
† Edser, *op. cit.*, p. 408.

a given wave be in the water, and how much will it lag behind the corresponding wave in the air?

7. With a grating mounted on a spectrometer, a blue line in the third-order spectrum is seen to coincide with the red line (6438 Å) of the second order. What is the wave length of the blue line?

8. With a grating spectrometer, the second-order green line (5218 Å) is seen to coincide with a third-order line of the ultraviolet. What is the wave length of the latter line?

9. In Gale's interferometer, two pieces of plane glass 10 cm long are separated at the bottom by a piece of paper 0.0236 mm thick. If the distance on the glass from the first to the twenty-first fringe is 2.5 cm, what is the wave length of the light used?

10. Show that in Young's experiment, if s is the distance between the two slits, D the distance from the plane of the slits to that of the interference bands, and d the distance between adjacent bands, the wave length of the light used is given by the formula: Sd/D.

CHAPTER XLIII

DOUBLE REFRACTION: POLARIZATION

572. Double refraction. In 1669, Bartholinus of Copenhagen discovered that when a ray of light passes through a crystal of Iceland spar, or calcite ($CaCO_3$), it generally emerges as two rays. One of these obeys within the crystal the ordinary laws of refraction, while the other does not. The former is therefore called the **ordinary ray** and the latter, the **extraordinary ray** (Fig. 557).

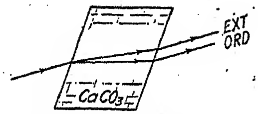


FIG. 557. Polarization by Double Refraction in Calcite

There is, however, one direction in the crystal, called the **optic axis,** along which double refraction does not take place. This axis makes equal angles with the edges of the crystal at the vertex where the obtuse angles meet. Substances like calcite, quartz, tourmaline, etc., which exhibit double refraction and have a single optic axis, are called **uniaxial** crystals. Others like mica, topaz, potassium nitrate, etc., which have two optic axes, are called **biaxial** crystals.

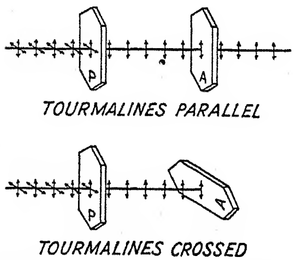


FIG. 558. Polarization by Absorption in Tourmaline

When tested by a radiometer, the ordinary and extraordinary rays are each found to have half the energy of the incident beam.

573. Polarization of light. The double refraction of tourmaline, discovered by Biot in 1816, is peculiar in that only the extraordinary ray emerges, the ordinary being absorbed within the substance itself. Biot found that if he looked at a source of light through a single crystal cut parallel to the optic axis, no change took place when the crystal was rotated. When he looked through two crystals, however, the light was transmitted freely if their axes were parallel, but was cut off almost completely if their axes were perpendicular to each other (Fig. 558).

Evidently the light that had passed through the first crystal possessed a two-sidedness that ordinary light did not have. Malus gave to such light the name **polarized light** (cf. Sec. 225).

The first plate of tourmaline is called the **polarizer *P***; the second one, which shows whether the light is polarized or not, is called the **analyzer *A***.

Tourmaline is usually green or red, and hence is not suitable for general purposes of polarization. It serves admirably, however, for testing the two rays transmitted by the rhomb of calcite. When these are observed through a tourmaline crystal, it is found that one ray is transmitted and the other extinguished for a certain position of the tourmaline; and that for a position at right angles to this, the first ray is extinguished and the other transmitted.

Hence **the ordinary and extraordinary rays produced by double refraction are polarized at right angles to each other.**

As was shown in Sec. 225, only transverse waves can be polarized. Thus **polarization is the proof that light waves are transverse.**

Fresnel, who first recognized that the phenomenon of polarization required light to be transverse, assumed that the vibrations transmitted by tourmaline are those parallel to the optic axis. Light waves are now known to be electromagnetic, and it will be shown in Sec. 578 that what we call vibrations are changes in the electric vector (the magnetic vector being always at right angles to the electric). Accordingly, tourmaline transmits the ray whose electric vector is parallel to the optic axis. Fresnel's guess was right.

574. The Nicol prism. The most satisfactory device for producing polarized light is a prism invented by William Nicol, a skillful lapidary of Edinburgh, about 1827. Its construction is shown in Fig. 559.

The end faces, which naturally make angles of 71°, respectively, with two opposite edges of a crystal of Iceland spar, are ground until this angle is 68° as shown. The crystal is then sawed in two, perpendicular to these end faces along the short diagonal as shown, and the sawed faces are polished. The resulting polyhedrons are then cemented back together with Canada balsam.

The index of refraction (for sodium light) is, for the ordinary

ray, 1.658; for the extraordinary ray, 1.486; and for the Canada balsam, 1.530. Hence the ordinary ray passes from an optically denser to an optically rarer medium; its vibrations are perpendicular to the paper (as shown by the dots); and it meets the surface of the balsam at an angle greater than the critical angle. It therefore suffers total reflection at the surface of the balsam and is totally absorbed by the black paint, as shown. The extraordinary ray, however, is passing from a rarer to a denser medium;

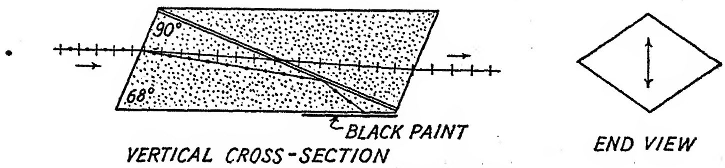


FIG. 559. The Nicol Prism

its vibrations are parallel to the plane of the paper (as shown by the short cross lines), and it re-enters the calcite at an angle less than the critical angle. Consequently it is not totally reflected, but emerges plane-polarized.

The experiments of Wiener, Drude, and others * on the production of stationary waves in plane-polarized light identify what we have called the direction of vibration with that of the electric vector of Maxwell's electromagnetic theory. It follows, therefore, that in a beam of plane-polarized light transmitted by a Nicol prism, the electric vector is parallel to the short diagonal of the end face of the prism.

The Nicol prism, or *nicol,* is widely used in polarizing instruments because the light in passing through it undergoes no change in color.

The Glan Thompson prism is similar to the nicol, but its surfaces are so cut that the axis of the transmitted beam suffers no displacement when the prism is rotated. This is necessary in such instruments as spectrophotometers.

575. Polaroid. In the last few years, a material for polarizing light has been brought out under the trade name of *polaroid.* This is made by suspending tiny synthetic crystals of iodized quinine sulphate in liquid cellophane, which is then hardened under conditions that cause practically all the crystals to arrange them-

* Edser, *Light for Students,* p. 484.

selves with their axes parallel. Each square inch of film contains about 10^{12} crystals, which have a slight purple color. Polarization takes place by double refraction followed by the absorption of one component, as in tourmaline, and is about 99.5% complete.

Polaroid film is mounted between two plates of annealed glass and set in a bakelite rim on which the axis is marked (Fig. 560). Polaroid plates of almost any size may be made, and they are particularly serviceable where large ones are required, as in apparatus for photo-elasticity. They are also being applied to such practical problems as the elimination of glare.

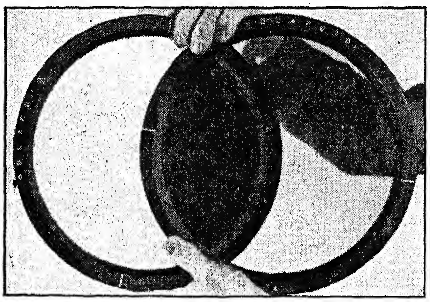


FIG. 560. Polaroids. (Courtesy Polaroid Corp., Boston)

576. Polarization by reflection and transmission. In the year 1808, Malus, a French military engineer, while looking through a crystal of calcite at the sun reflected from a window of the Luxembourg Palace, noticed that as the crystal was turned, first one of the two images disappeared and then the other. He had discovered that light is polarized by reflection.

Light is always partially polarized when it is reflected. The phenomenon is conveniently studied by means of Norremberg's apparatus (Fig. 561).

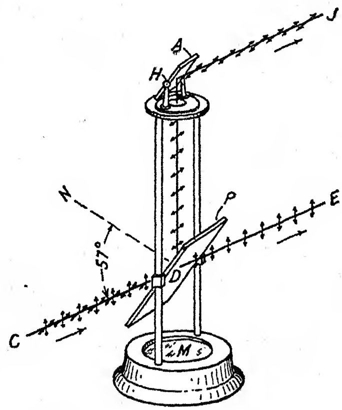


FIG. 561. Norremberg's Polariscope

Let a beam CD of ordinary light, with its vibrations in all directions at right angles to CD, fall upon a plate P of clear glass. Then let the reflected beam DH and the transmitted beam DE be examined with a Nicol prism, or a polaroid. It will be found that both beams are partially polarized at right angles to each other. The vibrations of the polarized part of the reflected beam are perpendicular to the

plane of incidence (i.e., parallel to the surface of the mirror); while the vibrations of the polarized part of the transmitted beam are in the plane of incidence, which is normal to the reflecting surface. In Fig. 561 only the polarized parts of the reflected and transmitted beams are shown.

The reflected light is said to be plane-polarized **in the plane of incidence,** which is called the **plane of polarization.** These definitions lead to the rather unfortunate convention that the vibrations (direction of the electric vector) are perpendicular to the plane in which the light is said to be polarized.

A second mirror A* serves as an analyzer. When parallel to P, it reflects readily, the beam HJ being polarized in the plane of incidence as shown. But when A is rotated 90° about DH as an axis, no light is reflected. The vibrations of DH are then normal to A and the beam is totally absorbed or transmitted, according as A is of black or clear glass.

As the angle of incidence CDN is increased, the degree of polarization of the reflected beam DH increases until, for an angle of about 57°, it is almost completely polarized, if the reflector is of crown glass. The degree of polarization of the transmitted beam is then about 16%. If the angle of incidence is then further increased, the polarization of the reflected beam again diminishes; but that of the transmitted beam continues to increase up to grazing incidence.

Since only about 9% of the energy of the incident beam is reflected at best, we cannot get a strong beam of polarized light by reflection. However, the degree of polarization of the transmitted beam increases at each reflecting surface, so that by passing the beam through a medium such as 8 microscope slides, a very strong beam about 60% polarized is secured.

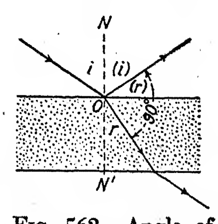


FIG. 562. Angle of Polarization

577. Brewster's law. Sir David Brewster discovered in 1815 that when maximum polarization takes place, the reflected and refracted beams are at right angles to each other. This fact enables us to compute the *angle of polarization*, i.e., the angle of incidence for which polarization is a maximum.

* For this mirror black glass is best because it gives no reflection from its back surface.

For the nonmetallic substance of which a mirror is made, we have (Fig. 562):

$$\begin{aligned}\mu &= \frac{\sin i}{\sin r}\\ &= \frac{\sin i}{\sin (90° - i)}\\ &= \frac{\sin i}{\cos i}\\ &= \tan i. \qquad (429)\end{aligned}$$

This is **Brewster's law: The tangent of the angle of maximum polarization equals the index of refraction of the reflector.**

As the index of refraction is different for different wave lengths, so is the angle of maximum polarization. Jamin found that for ordinary light, polarization is most complete when the index of refraction is 1.46.

Crown glass has a mean index of refraction of about 1.55; hence its angle of polarization is:

$$\begin{aligned}i &= \tan^{-1} 1.55\\ &= 57°10'.\end{aligned}$$

Brewster's law enables us to determine the index of refraction of opaque substances such as coal, felspar, obsidian, etc., but it cannot be used with metal reflectors since they produce elliptic polarization.

578. Polarization by scattering: The Tyndall effect. The phenomenon of the scattering of light by fine particles is known as the *Tyndall effect*, although it was first used by Faraday in 1857 to detect the presence of particles of gold which were too small to be seen by means of a microscope.

A familiar example of the effect is the path of a beam of sunlight admitted into a darkened room through a keyhole, or other aperture, and made visible by the light reflected from particles of dust in the air. A more spectacular illustration is the brilliant display of searchlights of the United States battle fleet at San Pedro, California (Fig. 563). Such beams, though not always conical, are called "Tyndall cones" because John Tyndall, who

first made a thorough study of the phenomenon, used a conical beam (1869).

He found that if a strong beam of light is passed through a vessel containing no fine particles, the beam is not visible. **But when fine particles such as those of smoke are admitted, the beam becomes visible;** and the scattered light is plane-polarized, with its

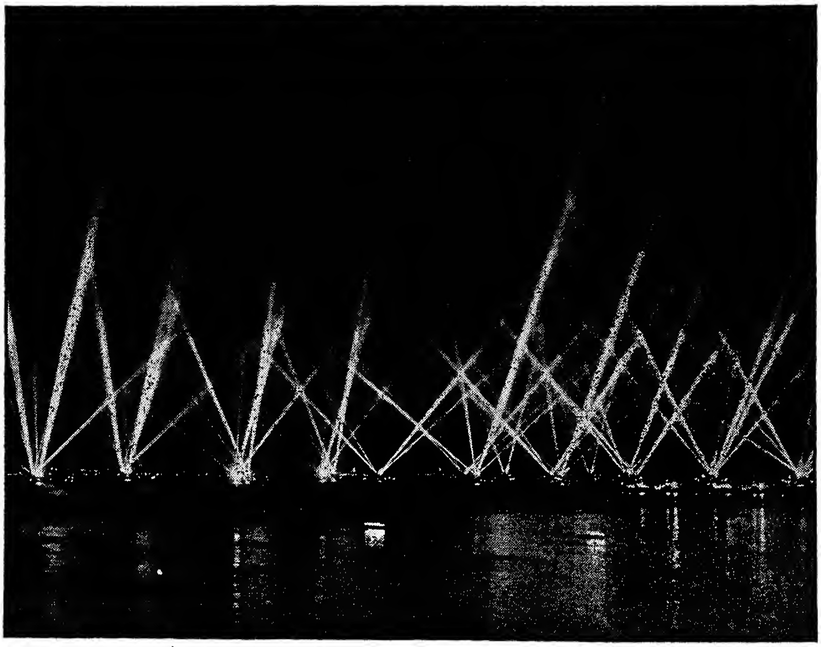


FIG. 563. Gigantic Tyndall Cones

vibrations perpendicular to the plane determined by the direction of the incident light and the line of observation (Fig. 564).

The color and the intensity of the scattered light depend upon the size of the particles. When they are very small (comparable to the wave length of the light), the scattered light is blue; **hence the sky is blue.** As the particles are made larger, the longer wave lengths also are scattered, until finally the scattered light is white. Thus the smoke from the lighted end of a cigar is blue, but that exhaled from the lungs is white on account of the grouping of the carbon particles into larger assemblies by the moisture of the breath. Ordinary reflection does not affect color, so that scatter-

ing by fine particles is due to diffraction. With large particles the light is reflected.

This phenomenon also enables us to determine the direction of the electric vector in a beam of plane-polarized light. The scattering is most probably due to sympathetic vibrations of electrons in the atoms with the electric vector of the original beam. In this beam all the vibrations are normal to its direction *SO*. Of these vibrations, only the components normal to the plane *SON* will be perceived by the eye *E* as light, since light waves must be transverse. In the experiment, the polarized beam *ON* is transmitted most completely by the nicol when the short diagonal of its end face is normal to the plane *SON*. Hence, in a beam of plane-polarized light transmitted by a nicol, the plane of the electric vector is parallel to the short diagonal of the end face of the nicol.

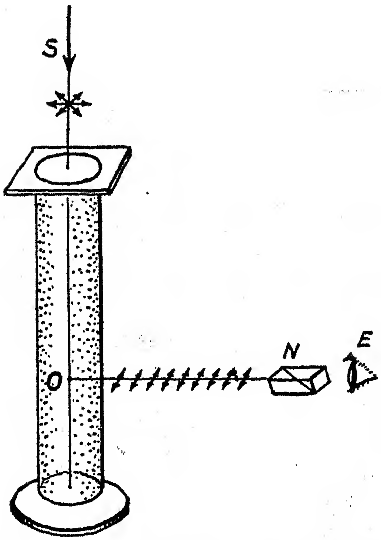


FIG. 564. Polarization by Scattering

579. Interference of plane-polarized light. In Young's experiment it was found that in order for two wave systems to interfere, they must originate from the same source. In addition, if the two beams are plane-polarized, the planes of polarization must be the same.

An arrangement for producing such interference is shown in Fig. 565.

A beam of ordinary white light *MN*, whose vibrations are transverse in all possible planes through *MN*, strikes the polarizer *P* and the emerging beam *NO* is plane-polarized with its vibrations (electric vector) parallel to the optic axis a_1a_1,* which is horizontal.

It then passes through a doubly refracting crystal *C*, which re-

* For maximum effect, a_2a_2 should make an angle of 45° with a_1a_1.

solves it into an ordinary ray OR (with its vibrations perpendicular to the optic axis a_2a_2) and an extraordinary ray XY (with its vibrations parallel to the optic axis). Assuming that the index of refraction for the extraordinary ray is less than for the ordinary,

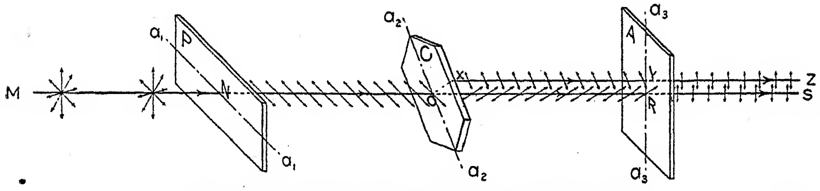


FIG. 565. Interference of Plane Polarized Light

it follows by Sec. 551 that the extraordinary ray travels through the crystal faster than does the ordinary. Hence, on emerging, there will be a phase difference between them; but since their vibrations are not in the same plane, they will not interfere.

If, however, an analyzer A is interposed in the path, with its axis at right angles to that of the polarizer, it will transmit only the vertical components of OR and XY in such a way as to introduce an additional 180° difference of phase, as may be seen in Fig. 566. Here point B represents the axis of Fig. 565 when viewed from M toward S; and a_1a_1 and a_3a_3 are the traces of the planes of vibration of the beams transmitted by the polarizer and analyzer, respectively, on the plane of the paper.

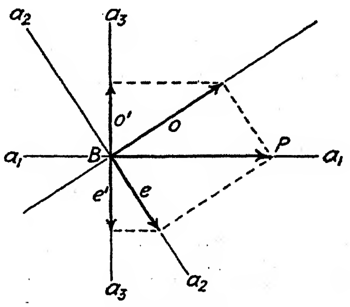


FIG. 566

The vector BP, representing the amplitude of the beam transmitted by the polarizer, is resolved by the crystal into the extraordinary beam e and the ordinary beam o. Of these beams, the analyzer transmits only the components e' and o', which are parallel to its axis, and introduces a phase difference of 180° due to the resolution, as shown. This is in addition to the phase difference due to their relative retardation by the crystal.

Beams RS and YZ (Fig. 565) are therefore in a condition to interfere; and they will be in opposite phase for any color for which their relative retardation in the crystal is an even number of half wave lengths. That color will then suffer destructive interference.

Hence a crystal which is colorless in ordinary light shows colors due to interference when examined in polarized light, the colors depending upon the index of refraction and the thickness of the specimen. This fact is employed by the geologist to detect differences in structure of crystals that otherwise escape notice.

580. Circularly polarized light. If monochromatic light is used in Fig. 565 and the thickness of the doubly refracting crystal C is such as to retard the ordinary ray one-quarter of a wave length behind the extraordinary, we have two simple harmonic vibrations at right angles and 90° out of phase. These are the conditions which give circular motion. Consequently, the beam between C and A is said under these circumstances to be **circularly polarized:** the plate C is then called a **quarter-wave plate.**

When circularly polarized light is passed though an analyzer A, the emerging beam is plane-polarized, but its intensity is not changed as the analyzer is rotated. If, however, circularly polarized light is passed through a quarter-wave plate, it again becomes plane-polarized and can then be extinguished by the analyzer A. This is the test for circularly polarized light.

If the ordinary and extraordinary beams are not 90° out of phase, on rotating the analyzer the intensity of the emerging plane-polarized beam will vary. The incident beam is then said to be **elliptically polarized.**

581. Photoelasticity. Sir David Brewster reported to the Royal Society in 1816 that glass and other transparent bodies under stress exhibit the property of double refraction. This phenomenon, known as *photoelasticity*, has come into extensive use by engineers for the determination of the stresses in machines and structures that do not admit of satisfactory solution by ordinary mathematical analysis.

A small model of the structure is made to scale from some transparent material, such as bakelite, and loaded in a manner similar to the loading of the structure under investigation.

If the loaded model is placed in ordinary plane-polarized light between crossed nicols and viewed through the analyzer, it is seen to be crossed by a system of colored curves (*isochromatics*) and a system of dark curves (*isoclinics*). The former are the result of interference, as described in Sec. 578. The latter are due

to the fact that where the direction of a principal stress in the model is parallel to the direction of vibration of the plane-polarized light, that light will pass through the model unaltered and will be stopped by the analyzer, thus giving rise to dark bands. When monochromatic light is used, both sets of curves are dark and confusion may arise. This may be avoided by using circularly polarized light, which eliminates the isoclinics. By means of these sets of curves and characteristic constants of the material, the stresses in the model may be determined, and these in turn are proportional to those in the actual structure.

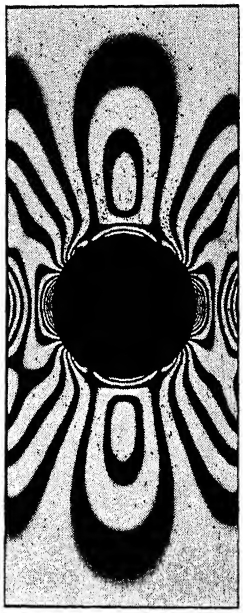


FIG. 567. Fringes Showing Stress Distribution around Punched Hole. (Courtesy of Dr. M. M. Frocht)

Figure 567 shows the isochromatics, or "fringes," for a test piece having a hole in it and subjected to pure tension. The curves are not stress contours, but are lines along which the difference of the principal stresses in the member is constant. At a surface, however, where one of the principal stresses is zero, the value of the remaining stress is proportional to the number of the corresponding fringe, counted from the point where the stress is known to be zero.

One may determine by inspection the region of greatest stress by the concentration of fringes. Thus, in Fig. 567, the stress is not uniform across the material on each side of the hole, but is greatest at the surface of the hole.

582. Rotation of plane of polarization. In 1811 Arago discovered that various solids, liquids, and vapors have the property of rotating the plane of polarization of a beam of light passed through them in certain directions. These are called **optically active** substances. If the rotation is clockwise as one looks along the beam toward the source of the light, the substance is said to be **dextrorotatory**; otherwise it is **levorotatory.**

There are about 30 solids and some 700 liquids and solutions that are known to be optically active. Of the solids, quartz is the most notable, some crystals producing right-handed, others

left-handed, rotation. All the optically active liquids and solutions are organic compounds. Sugar which in solution is dextrorotatory is called **dextrose**; that which is levorotatory, **levulose.**

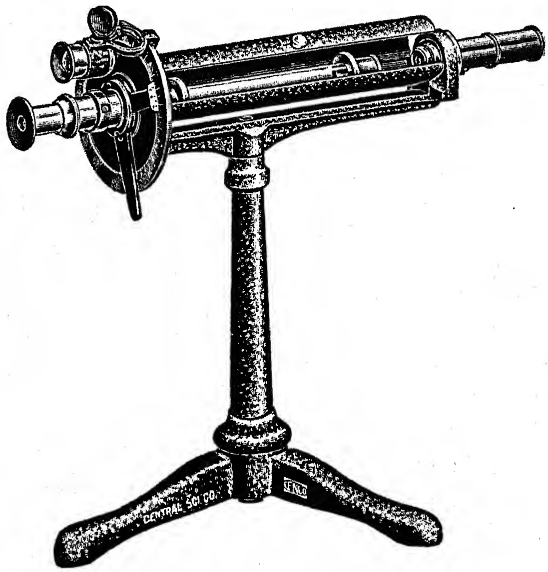


FIG. 568. Polarimeter. (Courtesy Central Scientific Co.)

Specific rotation is defined as the rotation produced by a column of liquid one decimeter long containing one gram of the optically active substance per cubic centimeter of the solution. It is positive for dextrorotatory substances and negative for those that are levorotatory.

Algebraically,

$$\alpha \equiv \frac{r}{lc} \tag{430}$$

where α is the specific rotation;
r is the angle of rotation in degrees;
l is the length of the column in decimeters; and
c is the density of the liquid or the concentration of the solution in grams per cm^3.

In Fig. 568 is shown a polarimeter for the measurement of optical activity.

583. The Faraday and Kerr effects. Faraday found (1845) that if a transparent isotropic substance is placed in a strong magnetic field, the plane of polarization of light transmitted through the medium parallel to the direction of the field is rotated. This is known as the **Faraday magneto-optical effect.** The direction of rotation is the same as the direction of the current that produces the magnetic field. If the beam of light is reflected back along the same path, the effect is doubled; whereas in the case of natural, optically active substances, the two rotations neutralize.

Faraday sought a similar effect due to an electrostatic field, but without success. It was not until 1875 that the phenomenon was discovered by Kerr. He found that certain substances ordinarily isotropic become doubly refracting when in an electrostatic field, and have the properties of a uniaxial crystal with its optic axis parallel to the direction of the field. This is known as the **Kerr electro-optical effect.** It is particularly marked in carbon bisulphide and nitrobenzene.

To demonstrate the phenomenon, a "Kerr cell" is prepared. This consists of a glass box with two metallic sides which form the plates of a condenser. The box is filled with the substance to be examined and a high potential is established between the metallic sides. The light is transmitted parallel to the condenser plates.

If the cell is placed between a pair of nicols set for extinction with their axes at 45° to the direction of the field, the arrangement forms an electrostatic shutter, transmitting when the field is on and stopping the light when the field is off. It opens and closes with extreme rapidity, and is therefore employed in investigations involving very short time intervals.

PROBLEMS

1. What angle must a beam of light make with a plate of flint glass ($\mu = 1.65$) for maximum polarization of the reflected beam?

2. With sodium light, if the index of refraction of Iceland spar is 1.658, and of Canada balsam, 1.530 for the ordinary ray, find the critical angle for Iceland spar with reference to Canada balsam.

3. If at maximum polarization 8% of a beam of light is reflected at the front surface of a plate of glass and 6.7% at the back surface, and if 6% is absorbed in the glass, what is the intensity of the emergent beam and its degree of polarization after passing through one plate?

4. Light reflected from the cleavage surface of a piece of coal is found to

experience maximum polarization for an angle of incidence of 57°30′. What is the index of refraction of the piece of coal?

5. A column of sugar solution 20 cm long, having a specific rotation of 66.4, rotates the plane of polarization 23.9°. What is the concentration of the solution?

6. A column of a solution of dextrose is 10 cm long and has a concentration of 0.12 gm/cm^3. If it rotates the plane of polarization 6.3°, what is the specific rotation of dextrose?

7. A solution of levulose has a concentration of 0.21 gm/cm^3, and in a column 20 cm long it rotates the plane of polarization $-37.2°$. What is the specific rotation of levulose?

CHAPTER XLIV

PHOTOMETRY

584. The standard candle. In order to assign numbers to the intensity, or strength, of a source of light in a given direction, it is necessary to adopt a standard source whose intensity is taken as unity.

The unit of luminous intensity is called the **candle.** Although originally defined in terms of candles made according to strict specifications, it no longer has any connection with an actual candle.

The *candle* is the intensity in a given direction of an imaginary source whose ability to give out light in that direction is equal to a definite fraction of the average horizontal intensity of a group of 45 carbon filament lamps preserved at the U.S. Bureau of Standards.* It is practically identical with the international candle established in 1909 by agreement among the national physical laboratories of France, Great Britain, and the United States. It is 1/10 of the intensity of the British Harcourt pentane lamp, and 10/9 that of the German Hefner lamp, to a close approximation.

An imaginary source whose luminous intensity in every direction is 1 candle is referred to as the *standard candle.*

A new definition of the candle has been adopted by the International Commission on Illumination, as follows:

The candle is 1/60 of the intensity in a given direction of one square centimeter of surface of a black-body radiator at the temperature of freezing platinum (1755°C), the surface being normal to the given direction. The new value of the candle is very slightly less than its former value.

The **candle power** I of a source in a given direction is its intensity in that direction expressed in candles. Thus, if the ability of a certain lamp to emit light in a certain direction is 25 times

* Illuminating Engineering Society, *Nomenclature and Photometric Standards* (1932).

that of the standard candle defined above, its candle power in that direction is said to be 25.

The intensity, or candle power, of a source in different directions varies greatly, as is shown by the polar diagram of Fig. 569. Here the length of the radius vector of any point of the curve represents the candle power of the lamp in that direction. The whole curve shows the distribution of intensity of the lamp in a vertical plane.

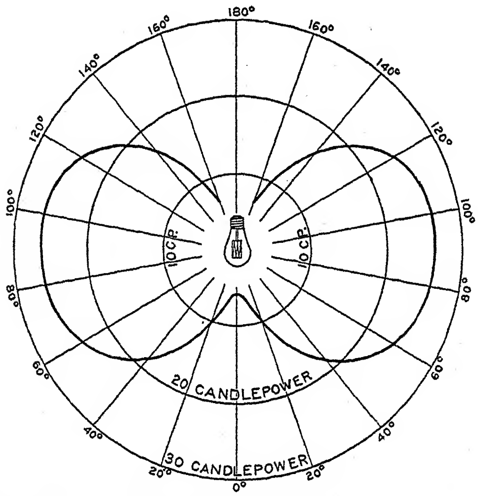


Fig. 569. Intensity Distribution in Vertical Plane

A **point source** is a source whose dimensions are negligible, so that its light may be considered to emanate from a point. If, in addition, its intensity is the same in all directions, it is a **uniform point source.**

In this chapter, unless otherwise specified, it is assumed that the source is so small in comparison with the other dimensions of the setup that it may be considered a point source.

585. Light: Luminous flux. *Light* is defined as radiant energy evaluated in proportion to the sensation of vision that it produces. Quantitatively, the amount of light of a given wave length is the amount of radiant energy of that wave length multiplied by the visibility factor (Sec. 594) for that wave length.

Luminous flux is the time rate of flow of light. It is a scalar quantity.

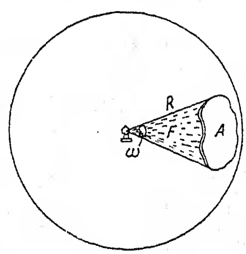


Fig. 570

Let us think of the light from a *uniform point source* of intensity I (Fig. 570) as falling upon an area A on the inner surface of a sphere of radius R having its center at the source.

Then the luminous flux F, or the quantity of light per second, falling on the area A will vary as follows:

$$
\begin{aligned}
F &\propto I && \text{when } A \text{ and } R \text{ are constant} \\
&\propto A && \text{when } I \text{ and } R \text{ are constant} \\
&\propto \frac{1}{R^2} && \text{when } I \text{ and } A \text{ are constant, by Eq. (194).}
\end{aligned}
$$

Therefore,

$$F \propto \frac{IA}{R^2} \quad \text{when all vary}$$

or

$$F = k\frac{IA}{R^2} \tag{a}$$

where k is a constant depending upon the choice of units.

The **solid angle** ω in steradians at the vertex of a pyramid or cone is defined as the ratio of the area A intercepted on a spherical surface, whose center is at the vertex of the pyramid or cone, to the square of the radius of the sphere. Thus, in Fig. 570,

$$\omega \equiv \frac{A}{R^2} \quad \text{steradians.}$$

Eq. (a) may therefore be written:

$$F = kI\omega. \tag{b}$$

As no unit of F has been chosen so far, we proceed to choose that unit so as to make $k = 1$.

The unit of I is 1 candle.
The unit of ω is 1 steradian.

We therefore define the unit of F, called the lumen, as follows: **The lumen is the luminous flux produced in a solid angle of one steradian by a uniform point source of one standard candle at the vertex of the angle.**

Writing into Eq. (b) the values in this definition, we have:

$$1 \text{ lumen} = k(1 \text{ candle})(1 \text{ steradian})$$

so that $k = 1$ for these units, as was intended.

Hence, when these units are used, k may be omitted, and Eq. (a) may be written:

$$F = I\frac{A}{R^2}. \tag{431}$$

In Eq. (431), the area A, being on the surface of the sphere, is everywhere normal to the direction of I. If an area A' is used *not* perpendicular to the direction of I, its resolved part A which is perpendicular to I is, from Fig. 571:

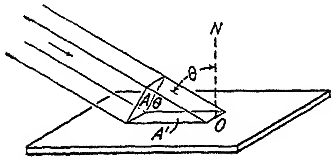


FIG. 571

$$A = A' \cos \theta.$$

Hence the more general expression for F is:

$$F = I\frac{A' \cos \theta}{R^2}. \tag{432}$$

In either case,

$$\frac{A}{R^2} = \frac{A' \cos \theta}{R^2} = \omega$$

so that always

$$F = I\omega \tag{433}$$

provided the luminous intensity I is uniform.

If I is not uniform, we must take a solid angle $d\omega$ so small that, within that angle, I may be considered uniform. Then Eq. (433) becomes:

$$dF = Id\omega. \tag{434}$$

The total solid angle ω_t about a point is the area of the sphere divided by the square of its radius. That is,

$$\omega_t = \frac{4\pi R^2}{R^2} = 4\pi \text{ steradians.}$$

The **total luminous flux F_t** from a uniform point source having an intensity of 1 candle in all directions is found by putting these values in Eq. (433).

$$F_t = \left(1 \frac{\text{lumen}}{\text{steradian}}\right)(4\pi \text{ steradians}) = 4\pi \text{ lumens.} \tag{435}$$

Since light travels in straight lines, the light that is radiated into a solid angle remains in that angle, however far its surface may be extended, so long as the medium is homogeneous and non-absorbing.

586. Illumination E is defined as the quantity of light that falls per second upon unit area of the illuminated surface; i.e., it is the luminous flux per unit of area A'. Algebraically,

$$E \equiv \frac{F}{A'} \tag{436}$$

so that, from Eq. (432),

$$E \equiv \frac{F}{A'} = \frac{I}{R^2}\cos\theta. \tag{437}$$

From this equation we get the units of illumination, for when

$$F = 1 \text{ lumen}; \quad I = 1 \text{ candle};$$
$$A' = 1 \text{ ft}^2; \quad R = 1 \text{ ft; and } \theta = 0;$$

then $E = 1$. This unit is called the *foot-candle*.

The foot-candle is the illumination on a surface all points of which are at a distance of one foot from a uniform point source of one candle (Fig. 572).

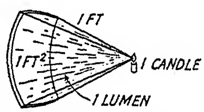


Fig. 572. The Lumen

Writing the above values back in Eq. (437),

$$1 \text{ foot-candle} \equiv 1\,\frac{\text{lumen}}{1 \text{ ft}^2} = \frac{1 \text{ candle}}{(1 \text{ ft}^2)} \times 1. \tag{438}$$

Similarly,

$$1 \text{ lux} \equiv 1\,\frac{\text{lumen}}{\text{meter}^2}$$

and

$$1 \text{ phot} \equiv 1\,\frac{\text{lumen}}{\text{cm}^2}.$$

587. Brightness. Whereas illumination refers to the light that falls upon a surface, *brightness* refers to the light that comes away from a surface. It may be measured in two ways: in terms of flux, or in terms of intensity.

1. The brightness B_f of a surface is defined as the luminous flux emitted or reflected per unit of **actual area** A' of the surface. This is often called the "flux brightness." If the brightness of the surface is not uniform, the definition gives its average brightness. Algebraically,

$$B_f \equiv \frac{F}{A'}. \tag{439}$$

When $F = 1$ lumen and $A' = 1$ cm², the unit of B_f is called the *lambert*.

$$1 \text{ lambert} \equiv 1 \frac{\text{lumen}}{\text{cm}^2}.$$

In words, **the lambert is the average brightness of any surface emitting or reflecting light at the rate of one lumen per square centimeter of actual surface.**

In the majority of cases, the brightness of a surface is not uniform and depends upon the direction from which it is observed. Consequently, the more generally used definition of brightness is the following.

2. The brightness B of a surface in a given direction is the ratio of its luminous intensity I in that direction to the **projected area** of the surface on a plane perpendicular to the given direction. Algebraically,

$$B \equiv \frac{I}{A} \equiv \frac{I}{A' \cos \theta} \tag{440}$$

where I is the intensity in the given direction;
A is the projected area;
A' is the actual area; and
θ is the angle between the direction of the intensity and that of the normal to the surface.

When I is in candles and A in cm², B is expressed in **candles per square centimeter of projected area.**

The relation between the candle per cm² of projected area and the lambert may be shown as follows (Fig. 573).

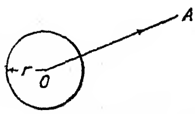


FIG. 573

Consider a sphere of radius r cm maintained at such a temperature that as a light source its intensity in any direction is 1 candle.

Then, in any direction OA, its intensity is 1 candle and its projected area is πr^2 cm². Hence its brightness is:

$$B = \frac{1}{\pi r^2} \frac{\text{candles}}{\text{cm}^2} \qquad \text{by Eq. (440).}$$

Again, the total luminous flux emitted is 4π lumens and the total area is $4\pi r^2$ cm². Hence, in terms of luminous flux,

$$B_f = \frac{4\pi \text{ lumens}}{4\pi r^2 \text{ cm}^2} = \frac{1}{r^2} \frac{\text{lumens}}{\text{cm}^2} \qquad \text{by Eq. (439).}$$

But these two values represent the same brightness. Therefore,

$$\frac{1}{\pi r^2}\frac{\text{candles}}{\text{cm}^2} = \frac{1}{r^2}\frac{\text{lumens}}{\text{cm}^2}$$

or

$$1\,\frac{\text{candle}}{\text{cm}^2} = \pi\,\frac{\text{lumens}}{\text{cm}^2} = \pi \text{ lamberts.}$$

The **reflection factor,** or **reflectivity,** of a surface is the ratio of the light reflected by the surface to the light incident upon it. In general, the reflection factor will be different for different wave lengths. Considering all wave lengths together, i.e., "white light," the reflection factor is called the **albedo** of the body.

The distinction between illumination and brightness may be made clearer by the following illustration. Let a piece of white blotting paper and a piece of black velvet be placed side by side at a distance of 2 ft from a 40-watt lamp (about 40 cp). The illumination is then the same for both paper and velvet: about 10 foot-candles, or lumens/ft^2.

But the albedo of white blotting paper is 70%, whereas that of black velvet is only 0.4%. Hence the brightness of the paper will be about 7 lumens per ft^2, but that of the velvet will be about 0.04 lumens/ft^2. Black letters on the paper could be easily read, but on the velvet they would be quite invisible.

In general, if

E is the illumination on a reflector in foot-candles,
ρ is the reflectivity of the surface, and
B is the brightness of the surface in candles/ft^2 in a direction normal to the surface,

then

$$B = \rho\frac{E}{\pi}. \tag{440a}$$

588. Lambert's cosine laws. In Eq. (437), it is seen that **the illumination of a surface is directly proportional to the cosine of the angle which the incident ray makes with the normal to the surface.** This fact is known as *Lambert's cosine law for incidence,* after Johann Heinrich Lambert.

Lambert concluded also, from experiment, that **the luminous intensity I emitted in a given direction by a perfectly diffusing surface is proportional to the cosine of the angle between the**

direction of the emitted ray and the normal. This is *Lambert's cosine law for emission.* It is rigorously true only for a theoretical black body, but is sufficiently accurate for practical purposes in many cases. It is in consequence of this law that a uniformly hot incandescent sphere appears equally bright at all points on its surface.

589. The Bunsen photometer. This is the standard device for comparing the luminous intensities (candle powers) of two sources.

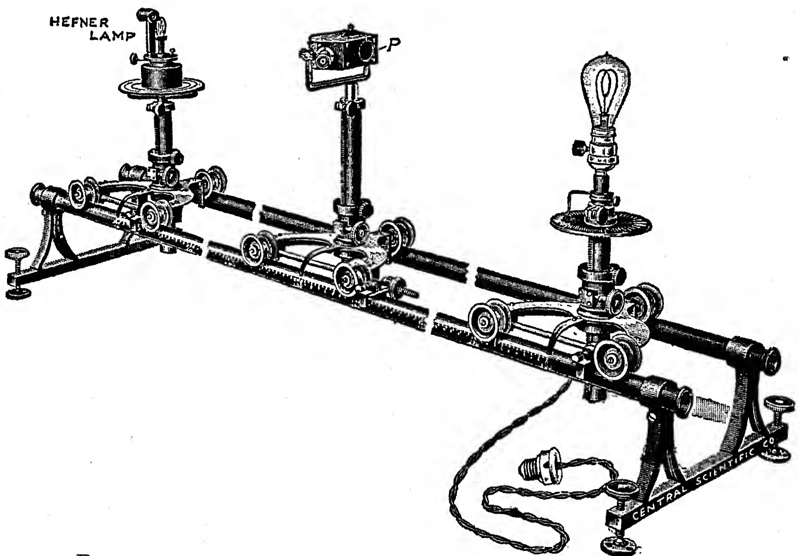


FIG. 574. Bunsen Photometer. (Courtesy Central Scientific Co.)

It consists (Fig. 574) of a sturdy steel bed, or **optical bench,** on which are three movable carriages. On the end carriages are mounted the two lamps to be compared, while on the middle carriage is placed the **photometer head,** or **photometer box.** Within this box is a screen whose two sides are illuminated by the respective lamps, and an arrangement of prisms or mirrors by means of which both sides of the screen may be seen side by side.

The screen is adjusted to a position between the two sources at which a **photometric match** is obtained; i.e., until the two sides of the screen appear equally bright. Then, if the two sides of the screen have the same reflection factor, the illuminations of the two sides are equal.

In Fig. 575, let I_1 and I_2 be the luminous intensities of the two

sources, respectively, and suppose the screen S to have been adjusted until a photometric match has been secured. That is,

$$E_1 = E_2$$

which, from Eq. (437), gives:

$$\frac{I_1}{D_1^2} \cos \theta_1 = \frac{I_2}{D_2^2} \cos \theta_2.$$

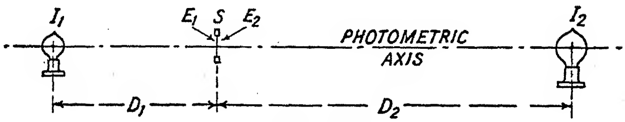


FIG. 575

But the instrument is made so that $\theta_1 = \theta_2 = 0$, since the beams from both lamps are normal to the screen. Hence,

$$\cos \theta_1 = \cos \theta_2$$

and

$$\frac{I_1}{I_2} = \frac{D_1^2}{D_2^2} \tag{441}$$

which is the **law of the Bunsen photometer.**

590. The Lummer-Brodhun cube. Of the various devices that have been designed for bringing the light from the two sides of the photometric screen into juxtaposition, the Lummer-Brodhun cube is one of the most satisfactory. It is shown at C (Fig. 576), in its position relative to the other parts of the photometric head.

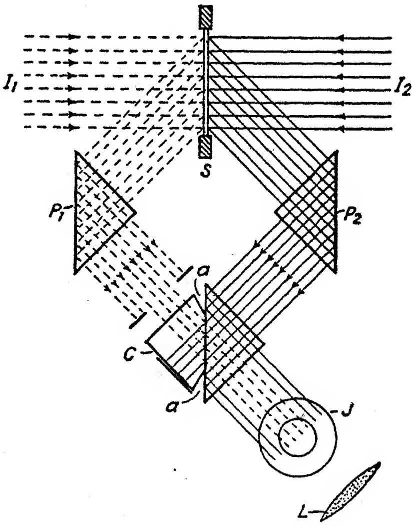


FIG. 576. Lummer-Brodhun Cube

The original cube of crown glass is sawed along a diagonal plane, yielding two triangular prisms. The corners of the hypotenusal face of the left prism are ground away, leaving that face elliptic in shape. The two sawed faces are then highly polished and cemented back together with

Canada balsam, whose index of refraction is very nearly the same as that of crown glass.

Light from the two sources I_1 and I_2 is received on opposite faces of the screen S, which is coated with the very white MgO or $MgCO_3$.

From S the beams are diffusely reflected into the total reflecting prisms P_1 and P_2, which in turn reflect them into the cube, as shown.

Rays (full lines) from I_2, which strike the hypotenusal face of the right-hand prism where it is backed by air in the angles a, a, are totally reflected and illuminate the outer ring of the field of view J. Those rays from I_2 which strike the surface where the prisms are cemented together continue straight through and are absorbed by the black paint on the outside of the prism.

Similarly, only those rays (dotted) from I_1 that strike the central, cemented section of the left prism pass through, all the others being reflected or refracted to the sides and absorbed by the black paint.

Hence, in the field of view J as seen through the lens L, the central circle is illuminated entirely by light from I_1, while the surrounding annular ring is illuminated entirely by I_2. Their brightnesses can therefore be matched with a considerable degree of exactness, provided the lights from the two sources are of the same color.

When the lights from the sources are not of the same color, it is not possible to say just what is a photometric match. For this case it is customary to use the combined brightness and contrast type of Lummer-Brodhun cube * or a flicker photometer.

591. The flicker photometer. A simple form of flicker photometer has been devised by Whitman (Fig. 577).

A disk R of the form shown in Fig. 577b is mounted on the shaft AB and driven by a variable speed motor. It is set so that its face F makes an angle of 45° with the photometric axis. The stationary screen S also makes 45° with this axis, and both are coated white with magnesium oxide. They are viewed through the sight tube T.

Light is reflected to the eye from source I_1 when sectors V_1 and

* Cady and Dates, *Illuminating Engineering* (New York, John Wiley and Sons, 1928), p. 195

V_2 are in front of the tube, and from source I_2 when the openings G_1 and G_2 are in front of it.

The eye will therefore observe a flicker both on account of the difference of color and the difference of intensity of the sources. As the speed of the disk is increased, the color flicker will disappear first. Intensity flicker is eliminated by adjusting the distances

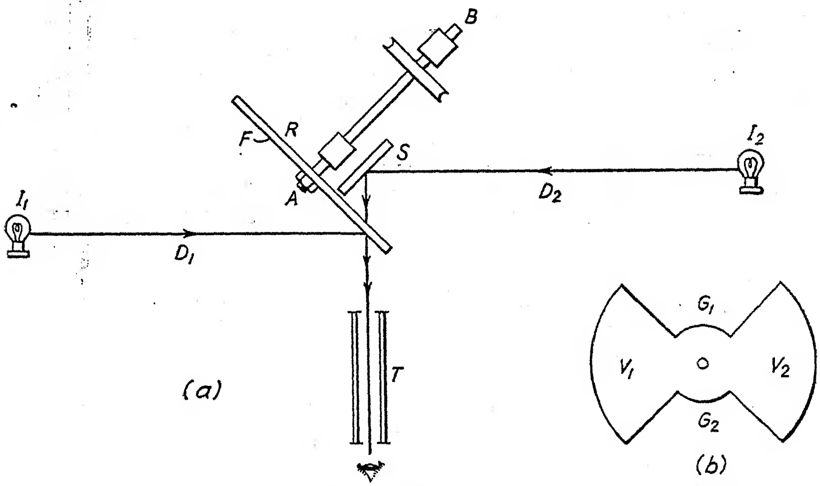


Fig. 577. Flicker Photometer

D_1 and D_2 of the sources from the center line of the sight tube. The best speed is the minimum one at which flicker can be made to disappear completely.

The intensities I_1 and I_2 are then directly proportional to the square of the distances D_1^2 and D_2^2 as in the case of the Bunsen photometer.

592. Illuminometers. Portable photometers, calibrated to read in foot-candles or other units of illumination, are called *illuminometers.* One of the most satisfactory of these has been the "Macbeth Illuminometer" of the Leeds and Northrup Company. It is shown in Fig. 578.

The Lummer-Brodhun cube is at C. The illumination E_2 to be measured is matched against the illumination E_1 on the translucent screen S produced by the lamp L. L is adjusted back and forth by a rack and pinion until a match is secured; the value of E_2 is then read directly on a scale on the rod R.

This instrument has the unique advantage that it is calibrated to the eye of the observer each time it is used, by comparison with an accompanying standard of illumination.

A very simple and inexpensive illuminometer, or foot-candle-

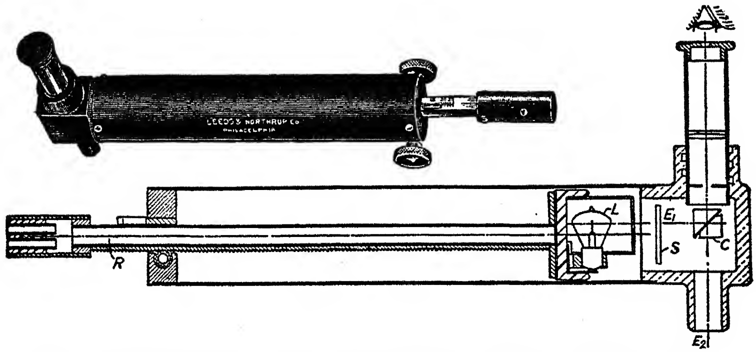


FIG. 578. Macbeth Illuminometer. (Courtesy Leeds and Northrup Co.)

meter, has recently been brought out. This employs the Weston photronic cell (Sec. 523), and a sensitive microvoltmeter (Fig. 579). While the instrument gives objective readings, it has the disadvantage that the sensitivity curve for the photronic cell is not the same as that for the eye, although it is a good approximation thereto. Compare Figs. 504 and 580.

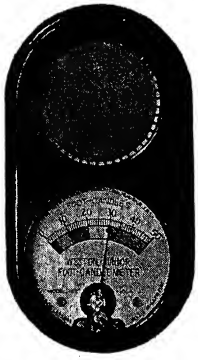


FIG. 579. Foot-Candle Meter. (Courtesy Western Elect. Inst. Co.)

With this instrument, Lambert's cosine law may be readily demonstrated.

593. The mechanical equivalent of light. The total energy transmitted per second in a beam of light may be measured by allowing the light to impinge upon the blackened front surface of a disk of silver, or gold, to whose back surface is soldered a sensitive thermocouple (Fig. 263). The reading of the thermocouple is proportional to the energy received (and converted into heat) per second, i.e., to the mechanical power of the beam in watts.

We may measure the quantity of light transmitted per second in the beam by matching it against the known luminous flux from a standard lamp on a photometer.

The *mechanical equivalent* of light m is defined as the number of

watts of radiant power Φ required to produce 1 lumen of luminous flux F.

Algebraically,

$$m \equiv \frac{\Phi}{F} \frac{\text{watts}}{\text{lumen}}$$

or

$$\Phi = mF.^*$$

Unfortunately it is not quite so simple, because the power required to produce 1 lumen is different for different wave lengths, being greatest in the violet and red and least in the greenish yellow. Hence it is necessary to specify the wave length λ by a subscript:

$$\Phi_\lambda = m_\lambda F_\lambda. \tag{442}$$

The mechanical equivalent of light has been determined by various observers, and is found to be a minimum for the wave length 5576 Å. The value of the

$$\text{Least mechanical equivalent of light} = 0.00162 \frac{\text{watt}}{\text{lumen}}.\dagger$$

594. Visibility. The **visibility factor** K_λ is defined as the number of lumens produced by 1 watt of radiant power; i.e.,

$$K_\lambda \equiv \frac{F_\lambda}{\Phi_\lambda}. \tag{443}$$

The visibility factor is therefore the reciprocal of the mechanical equivalent of light for a given wave length. Hence the visibility factor is a maximum for the same wave length (5576 Å) for which the mechanical equivalent is a minimum.

$$\text{Maximum visibility factor} = \frac{1}{0.00162} = 617 \frac{\text{lumens}}{\text{watt}}.$$

The **relative visibility** for any wave length is defined as the ratio of the visibility factor for that wave length to the maximum visibility factor. Relative visibility is usually expressed in percentage.

When relative visibilities are plotted as ordinates against wave

* Compare with mechanical equivalent of heat J: $W = JH$ (Sec. 317).

† W. W. Coblentz, *Bulletin*, U.S. Bureau of Standards, Vol. XIV, No. 3.

lengths as abscissas, the resulting curve (Fig. 580) is called the **visibility curve,** or the curve of retinal sensitivity.

The colors corresponding to the various wave lengths may easily be kept in mind by means of the following table of approximate values.*

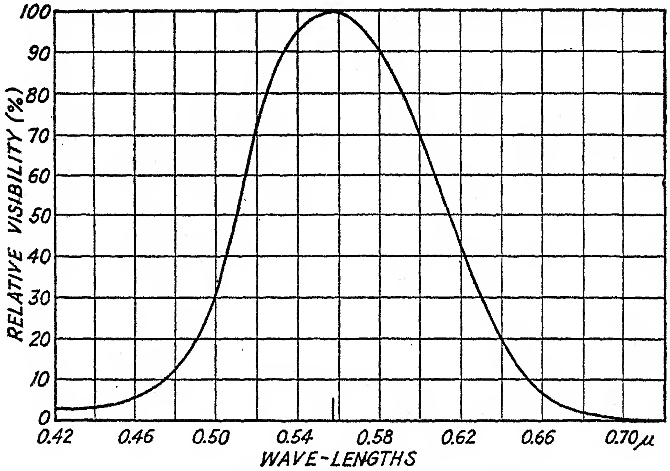


FIG. 580. Sensitivity of Human Eye

Colors	*Wave lengths*
Violet	$0.43\mu = 4300$ Å
Blue	$0.48\mu = 4800$ Å
Green	$0.53\mu = 5300$ Å
Yellow	$0.58\mu = 5800$ Å
Orange	$0.63\mu = 6300$ Å
Red	$0.68\mu = 6800$ Å

Here $1\mu = 0.001$ mm and 1 Å (angstrom) $= 10^{-8}$ cm.

It will be noticed that the wave length (5576 Å) for which visibility is a maximum lies in the greenish yellow, about midway of a normal spectrum. **It is very nearly the color of the light of the firefly.** The firefly is the most efficient light source known.

595. Quantity of light. We are now equipped to understand better the meaning of quantity of light, or simply of light, as the word "quantity" is often omitted.

* Michelson, *Wave Lengths and Their Uses* (University of Chicago Press, 1907), p. 17.

Since we have used the symbol Φ_λ for radiant power, it follows that

$$\text{Radiant energy} = \Phi_\lambda t.$$

Then if we use Q to designate quantity of light, we shall have, by the definition of Sec. 585,

$$Q_\lambda = K_\lambda(\Phi_\lambda t).$$

But by Eq. (443),

$$K_\lambda \Phi_\lambda = F_\lambda.$$

Therefore,

$$Q = Ft \tag{444}$$

whether the light is monochromatic or not.

When $F = 1$ lumen and $t = 1$ sec, $Q = 1$. This *unit quantity* of light is called the *lumen-second.*

A lumen-second is the quantity of light that passes a cross section of a beam of light in one second when the flux is one lumen.

PROBLEMS

1. A photometric match is obtained when the screen is 50 cm from a standard lamp of 40 cp and 120 cm from a tungsten lamp of unknown candle power. What is the candle power of the tungsten lamp?

2. A photometric match is obtained when the screen is 30 cm from a standard Hefner lamp and 150 cm from a carbon lamp. What is the candle power of the carbon lamp?

3. A photometric match is obtained when the screen is 40 cm from a standard lamp of 50 cp and 120 cm from a Nernst filament lamp. Find the candle power of the Nernst lamp.

4. A 20-cp and a 40-cp lamp are 100 cm apart on a photometric bench. Where must a screen be placed between them in order that the illumination on its two sides may be equal?

5. Lamps of 100 cp and 75 cp, respectively, are 200 cm apart on an optical bench. Where must a screen be placed in order that it shall be equally illuminated by the two lamps? Explain the two answers.

6. What is the total luminous flux from a 60-cp lamp?

7. What is the total luminous flux from a 75-watt lamp whose luminous efficiency is 0.8 watt per candle?

8. What is the illumination in lux at a point 6 ft from a 50-cp lamp?

9. The bulb of a table lamp is 2 ft from the printed page. What must be the candle power of the lamp to give the proper illumination for reading?

10. A 75-cp lamp is 4 ft above and 3 ft to one side of the center of a horizontal table top. What is the illumination at the center of the table top?

11. A 100-cp lamp is 6 ft above and 4 ft to one side of the center of a horizontal table top. What is the illumination at the center of the table top?

12. The top of a table makes an angle of 40° with the direction in which a lamp 6 ft away has an intensity of 60 cp. What is the illumination on the table at the corresponding point in foot-candles and in lux?

13. What is the quantity of light per sec that passes through an aperture of 1 dm^2, the center of which is 3 m from a 60-cp lamp and which makes an angle of 60° with the ray to its center?

14. Assuming the mirror of a headlight to be 2 ft in diameter and parabolic and to reflect 50% of the light from a 2000-cp lamp at its focus in a parallel beam, find the average intensity of the beam in lumens/ft^2.

15. Find the total illumination produced by the above headlight at a distance of 100 ft, due to both the direct and the reflected light.

16. The top of a desk is covered with blotting paper which has an albedo of 0.06. A 200-cp lamp is 4 ft above the table. What is the illumination at a point where the rays make an angle of 60° with the table top? Assuming perfect diffusion in any direction making an angle of less than 45° with the normal, what is its brightness in such a direction expressed in millilamberts and in candles/cm^2?

17. At a point 100 ft from the foot of a street lamp 10 ft high, the reflection factor is 0.25. The candle power in the direction of the point is 300. Assuming perfect diffusion when viewed at an angle of 20° with the normal, what is the brightness of the pavement in lumens/ft^2 and in millilamberts?

CHAPTER XLV

COLOR — SPECTRA

596. Newton's experiment. Most of us have noticed that when sunlight passes through the angles of crystal pendants or other pieces of "cut glass," little patches of rainbow colors are thrown upon the walls. In 1666, Sir Isaac Newton made a classic experiment at Trinity College, Cambridge, England, which reproduced this phenomenon on a large scale and which was the beginning of the scientific study of color.

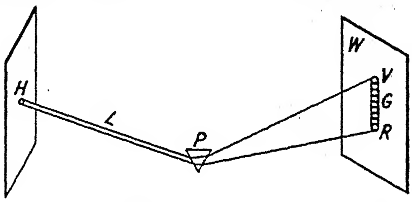


FIG. 581. Newton's Discovery of the Spectrum

Newton admitted a beam of sunlight L into a darkened room through a hole H in a window shutter and allowed it to pass through a prism P, as shown in Fig. 581. In so doing the beam of light was spread out into a band of colors on the opposite wall W. Such a band of colors he called a **spectrum.**

In this spectrum of sunlight he thought he could distinguish seven colors: violet, indigo, blue, green, yellow, orange, and red.

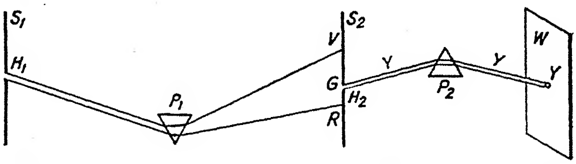


FIG. 582. A Pure Color Cannot Be Further Decomposed

(The initials of these colors form a mnemonic word, **vibgyor,** by which they are easily remembered.) Few, if any, observers can identify indigo as a distinct color.

Newton found also that the colors of the solar spectrum could be recombined into white * light by passing the beam through

* Of various proposed definitions of white light, none has yet been generally accepted. In this text we shall consider white light to be **average noon sunlight** reflected from light northern clouds.

two additional prisms properly placed; but that a narrow band of color from the spectrum could not be further decomposed by passing through another prism (Fig. 582).

597. Color defined. We now know that instead of only seven colors, there is an indefinite number of colors in a spectrum, one for each wave length, which may be readily shown by Young's experiment or by means of a grating.

Color is the visual sensation that depends upon the wave length (or the frequency) of the light that enters the eye. It corresponds to tone in sound.

Strictly speaking, the study of color falls in the field of psychology, but it is advantageous to take it up briefly here. Michelson's table of Sec. 594 enables us to associate colors and wave lengths in a rough way.

598. Rainbows. From Newton's experiment, we are able to explain the formation of rainbows. In that experiment we saw that the shorter wave lengths are deviated more than the longer ones, as was noted in Sec. 554. It often happens that rain is falling when the sun is bright and behind the observer. If at the same time the angle of elevation of the sun is correct, rainbows may be seen.

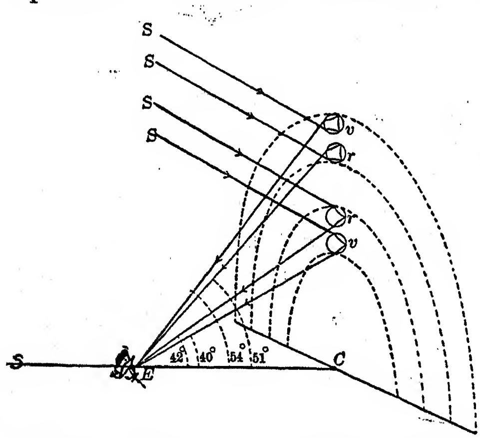


FIG. 583. Primary and Secondary Rainbows

Each raindrop acts as a tiny prism (Fig. 583), refracting the light when it enters and again when it leaves. Within some drops the light is reflected once; in others, twice. The former gives rise to the **primary bow**, and the latter to the **secondary bow.** It may be shown that the rays corresponding to the different colors are those for which the deviation is a minimum for their respective colors.

Since the sun is so far away that its rays are practically parallel, it will be clear from Fig. 583 that the drops which are in correct positions to direct light of a given color into the observer's eye will

lie on the surface of a cone whose axis is the line through the eye E and the center of the sun, and that the altitude of the sun must be within certain limits or a bow cannot be seen.

Since violet is refracted more than red, *in the primary bow* red will be seen on the outside of the bow, i.e., from drops higher up; and violet will be seen on the inside of the arch, i.e., from drops lower down.

In the **secondary bow** the light enters the drops near the bottom and suffers two internal reflections. Energy is lost at each of these reflections, and hence the secondary bow is **dimmer than the primary**: it lies above the primary, and has the colors reversed.

Since the axis of the cones on which the various colors are seen is determined by the sun and **the eye of the observer,** it follows that each observer sees a different rainbow.

599. Hue and saturation. Light of a single wave length is called a **hue,** a **pure color,** or **monochromatic light.** Colors as observed in daily affairs are frequently mixtures of one or more hues with a certain amount of white light.

"The **saturation** of a color is its degree of freedom from admixture with white light." * Thus monochromatic light is 100% saturated.

Any color may be specified in terms of its hues and their saturations. Pink, for example, might have the hue (λ = 6438 Å) red, and be 30% saturated; i.e., this pure red is mixed with 70% of white light.

600. Color of bodies. The colors of bodies in nature are never pure, being always the synthetic effect of many wave lengths. Only by taking a narrow strip from a spectrum do we get a color that is even approximately pure, or **monochromatic**; i.e., all of a single wave length.

When light falls upon a body there are only three things that can happen to it:

1. It may be absorbed.
2. It may be reflected.
3. It may be transmitted.

* "Nomenclature and Photometric Standards," Bulletin of the I.E.S. (Dec. 19, 1932).

Actually, a part of its energy may be dispersed in each of these three ways.

Most bodies exhibit the property of *selective absorption* or *selective reflection*, or both; that is, they absorb or reflect only certain wave lengths, which they thus select. For example, if a piece of red glass is held in a beam of white light, we call the emergent light red because red predominates. But if that light is then passed through a spectroscope, it will be found to consist of some orange and perhaps other colors besides red. Similarly, blue cobalt glass actually transmits violet, blue, and some orange and red. By blue light the red glass is black.

In the case of most transparent bodies, such as glass, the amount of light reflected is small. Hence **the light transmitted by bodies is that which they do not selectively absorb.**

This selective absorption is believed to be an effect of resonance. It is thought that within the body are particles whose natural frequencies of vibration are the same as the natural frequencies of certain colors in the incident light. These particles are accordingly set in vibration by resonance, taking up energy of the light whose frequencies are the same as their own, thereby converting it into molecular energy which is heat. The remaining frequencies are transmitted.

The color of many bodies which produce diffuse reflection is also the result of selective absorption. When a piece of red cloth or a red rose is held in white light, the light penetrates a very short distance into the body. Though this distance is extremely minute, it is sufficient for selective absorption to take place; and the light which is diffusely reflected is that which is not selectively absorbed. Bodies of this kind, if sufficiently thin, will appear the same color by transmitted light as by reflected light.

If a red cloth or rose is held in the blue end of the spectrum, it will appear black; for it absorbs blue, and there is no red present to be reflected.

Hence, **the apparent color of both transparent and ordinary opaque bodies depends upon selective absorption and the color of the incident light.**

Metals, on the other hand, owe their nominal colors to selective reflection. If white light impinges on a polished gold surface, the reflected light is yellow; whereas if a piece of yellow glass is used

as the reflector, the reflected light is still white. A glass surface reflects all wave lengths equally, but gold and other metals reflect certain wave lengths more strongly than others. This is the property which we call selective reflection. That this is different from the case of the red cloth is shown by placing a thin sheet of gold leaf in a strong beam of white light. The transmitted light is green, not yellow. Not all the yellow is reflected; some is absorbed.

Aluminum has very little selective absorption, hence it is used for surfacing astronomical mirrors. Various nonmetallic substances, e.g., aniline inks, potassium permanganate, and certain highly colored insects, exhibit selective reflection.

A **white** body reflects all colors in the same proportion: it appears white in white light, blue in blue light, etc. A **gray** is a weak white; i.e., a white of low intensity. The white walls of a room appear gray at twilight when they are feebly illuminated. A **black** body absorbs light of all colors, reflecting none.

601. Color of the sky. It will be recalled from Sec. 578 that Tyndall found the color of the light scattered by fine particles to depend upon the the size of the particles.

Lord Rayleigh has shown that the blue color of the clear sky is due to the light scattered by the molecules of the air itself, and that dust is not necessary, as was thought by Tyndall.* The gray, which is weak white, of an overcast sky is due to the light scattered by the large particles of water then suspended in the atmosphere. If there were no atmosphere, the sky would be black at all times as it is at night, and stars would be visible in the daytime.

The bright red and orange colors at sunrise and sunset are due to the transmitted light. We are then looking through a very thick layer of dust-laden air. The shorter waves of the blue end of the spectrum are scattered at right angles to the line of sight, so that in the light that gets through to the eye the longer waves of the red end of the spectrum predominate. In 1883, there was a great eruption of the volcanic island of Krakotoa which destroyed the entire island. Dust from this explosion encircled the earth for many days, and the colors at sunrise and sunset were exceptionally red during the period.

* Monk, *Light* (New York, McGraw-Hill Book Company, 1937), p. 286.

602. Color mixtures. Mixing colored lights and mixing colored pigments are very different operations and should be carefully distinguished.

Mixing colored lights is a process of addition. Thus, if we project upon a screen a circle of blue light, and from another source a circle of yellow light overlapping the blue circle, the region where they overlap will reflect to the eye both blue light and yellow light, and its apparent color will be the combined effect, or sum, of these two sensations perceived at the same time. That color will be white, if the wave lengths and intensities of the blue and yellow are properly chosen.

Mixing blue light and yellow light produces the sensation of white.

On the other hand, **mixing a blue pigment and a yellow pigment gives green,** which is accounted for as follows. Pigments, like lights, are never pure colors. A blue pigment absorbs red and yellow and reflects blue and green; a yellow pigment absorbs blue and violet and reflects yellow, green, and red. If white light falls upon a mixture of these pigments, we have:

White light contains	Violet	Blue	Green	Yellow	Orange	Red
Blue pigment absorbs				Yellow	Orange	Red
Yellow pigment absorbs	Violet	Blue				
Both reflect only			Green			

Hence **the mixing of pigments is a process of subtraction:** the light reflected is the color that is left after each pigment has subtracted the colors which it absorbs.

603. Complementary colors. Any two colors which together produce the sensation of white are said to be complementary. Helmholtz prepared a table of complementary colors from which some values are the following:

Red	(6562 Å)	and	Greenish-blue	(4921 Å)
Yellow	(5739 Å)	and	Blue	(4821 Å)
Violet	(4330 Å)	and	Greenish-yellow	(5636 Å)

The mixing is conveniently accomplished by pasting sectors of paper of the required colors on a disk of cardboard and whirling

it rapidly on the rotator (Fig. 584). If the colors are accurately chosen (which is difficult), the effect is gray, i.e., a weak white. Persistence of vision (1/10 sec) causes the sensation of one color to remain until another falls upon it, so that the combined effect is secured.

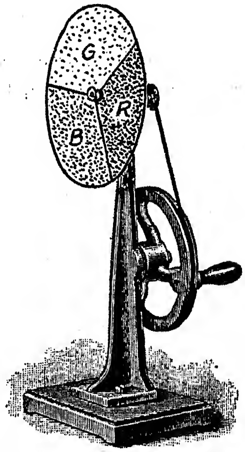


Fig. 584. Color Mixing Disk

604. Primary colors and primary pigments. While there is an infinite number of wave lengths, or pure colors, in the visible spectrum, it is found that only three—blue, green, and red—are necessary. By mixing these three in proper proportions all the color sensations of the spectrum may be produced synthetically.* Hence blue (4700 Å), green (5050 Å), and red (6710 Å) are called the **primary colors.**

In the process of color printing, however, a printer uses pigments, not colored lights. The **three primary pigments** used by printers are the complements of the three primary colors, viz., yellow, magenta, and peacock blue (a blue-green).

604a. Maxwell's color triangle. Maxwell and Helmholtz devised a color triangle (Fig. 565) which is helpful in the mixing of colored lights. It is an equilateral triangle at the corners of which are placed the primary colors, blue, green, and red. If any two of these are mixed in equal proportions, the resulting color is shown at the mid-point of the corresponding side of the triangle. If mixed in unequal proportions, the resulting color is represented by the point on the line joining the two colors, and dividing that

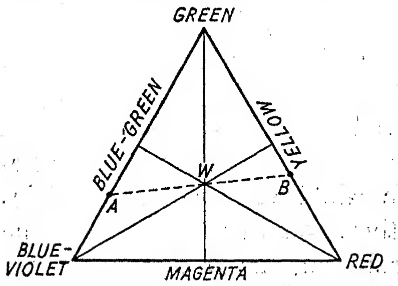


Fig. 585. Maxwell's Color Triangle

* Theories of color vision belong properly in the domain of the psychologists. Of the two theories which have received most attention, the rather recent one of Dr. Christine Ladd-Franklin seems to be generally preferred to the older theory of Young and Helmholtz.

line in the ratio in which the two colors are mixed, being nearer to the dominant color.

White is obtained by mixing the three primary colors in equal proportions, and is therefore represented by the center of the triangle. If any line AB is drawn through the center of the triangle, the colors represented by its extremities are complementary.

605. Fluorescence and phosphorescence. Various substances have the property of absorbing radiation of certain wave lengths and emitting the energy again as radiation of longer wave lengths. This phenomenon is called *fluorescence* or *phosphorescence* according as it ceases at once, or continues for a considerable time, after the exciting radiation is cut off.

Fluorescence is a notable property of fluorite (CaF_2), from which it takes its name. Just below red heat, fluorite becomes phosphorescent. Sulphate of quinine fluoresces blue when irradiated with ultraviolet light, whose waves are too short for it to be visible. Similarly, the fluoroscope, used in making x-ray examinations of internal organs of the body, consists of a glass screen coated with platino-barium-cyanide, which fluoresces a greenish-yellow under the action of x-rays whose wave lengths are of the order of .0001 the wave length of visible light.

Phosphorescence takes its name from phosphorus, but this is a case of mistaken identity. Phosphorescence is entirely different from the glowing of phosphorus, which is due to slow oxidation. Calcium sulphide shows phosphorescence for several hours, and the sulphides of barium and strontium show it even longer.

No satisfactory explanation of these phenomena has been found.

606. Color photography. In 1861 James Clerk Maxwell gave a lecture at the Royal Institution in London that marked the beginning of color photography. He exhibited three lantern slides (positives) made from three negatives he had taken of the same object: the first through a blue solution, the second through a green solution, and the third through a red solution. Using three lanterns, he projected these positives on to a screen through the same blue, green, and red solutions respectively that were used in making the negatives; and thus produced a picture of the object in its natural colors.

This illustrates the *additive process* of color photography, the

three primary colored lights being added on the screen. The method is cumbersome and obviously unsuited for the colored pictures that now appear in magazines.

These are produced by the following *subtractive process.* Three negatives are made of the object by means of a camera containing two partially coated mirrors which divide the beam of light coming from the lens into three beams. These beams are focused on three photographic plates: the first through a *blue-violet,* the second through a *green,* and the third through a *red* glass filter. From the negatives so obtained, three ordinary printer's cuts are made. The cut from the negative made with the blue-violet screen will have the regions depressed that were blue-violet in the original object, while the regions that had no blue-violet will be in relief, i.e. slightly raised above the surrounding surface. If this cut is inked with the color complementary to blue-violet, which is *yellow,* the print so made will reflect no blue-violet to the eye because that color is subtracted (absorbed) by the yellow. Similarly the cuts from the negatives made with the green and red filters will be inked with colors complementary to green and red, i.e. with magenta and blue-green respectively. These pigments in turn will absorb the green and red in the regions where these colors should not occur and will reflect all the other colors that do occur, so that all together a picture in proper colors is produced.

The *technicolor process* for moving pictures is similar to the above. The camera divides the beam of light from the lens into two beams. One of these passes through a green glass filter and produces a negative on panchromatic film that registers the greens of the scene photographed. The other beam passes through a magenta filter to two superimposed films, the first of which is sensitive to blue-violet light but not to red. On the back of this film is a dye that absorbs blue but not red, and the red is accordingly registered on the second film which is panchromatic (Fig. 585a).

From these three negatives, positive films are made which are so treated that parts which did not have green, blue-violet, and red respectively are slightly raised as in the case of the printer's cuts mentioned above. These positives are inked with colors complementary to green, blue-violet, and red respectively, and are printed in proper register on a single clear film which is used in the projector.

Consider a point on the final film corresponding to a place in the original scene that had no green. On the first positive, this region would be elevated and inked with magenta-colored ink.

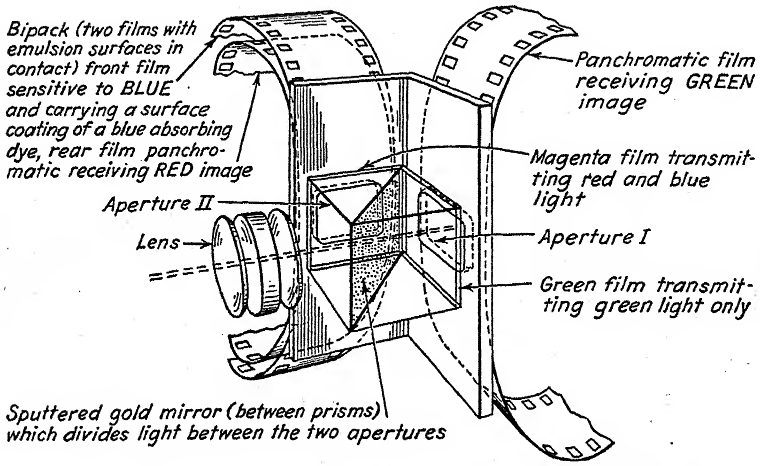


FIG. 585a. Technicolor Camera. (From Mees, *Photography*)

This ink would subtract (absorb) the green at this point from the white light coming from the source but would allow all the other colors at that point to pass through. The other two inks would similarly remove the blue-violet and the red where they should not occur; and the composite picture on the screen is therefore rendered in its proper colors.

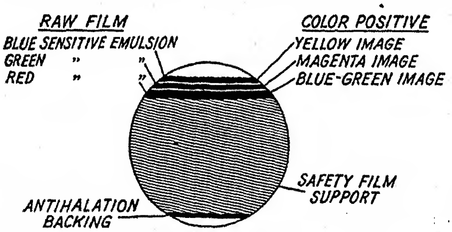


FIG. 585b. Kodachrome Film. (From Mees, *Photography*)

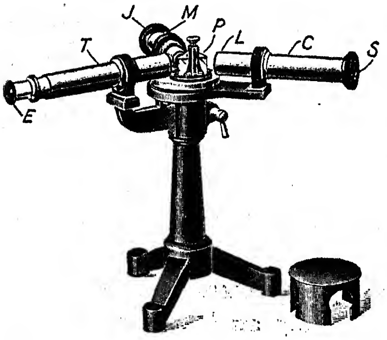
The *kodachrome* process of the Eastman Kodak Company is likewise a subtractive process. The film used in this process is built up in five layers (Fig. 585b). On top is a blue-sensitive emulsion containing a yellow dye, below this is a green-sensitive emulsion, then a layer containing a red dye to serve as a filter, and finally a red-sensitive emulsion next to the transparent supporting film which is coated on its other side with a non-halation backing.

When this quintuple film is exposed in an ordinary camera, the blue-violet component of the light forms an image in the first

layer, the green component in the second layer, and the red component in the fourth layer.

By a complex method of processing, the image formed in the blue-violet-sensitive layer is converted into a yellow positive; the image in the green-sensitive layer into a magenta positive; and that in the red-sensitive layer into a blue-green positive. The resulting triple positive when viewed by transmitted light or when projected on a screen yields a picture in the true colors of the original.*

607. The spectroscope. On account of the overlapping of the colors, the spectrum as obtained by Newton was far from pure.



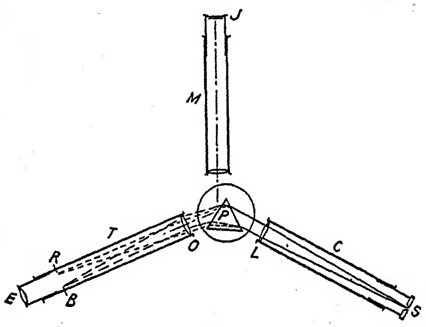


FIG. 586. Spectroscope

A **pure spectrum** is one in which the color at any point is due to a single wave length, which clearly is desirable. Obviously, the narrower the opening through which the light is admitted, the purer the spectrum will be. In a **spectroscope** (Fig. 586), this and other requirements are fulfilled with precision sufficient for most purposes.

The light to be examined is admitted through a narrow slit S whose sides are parallel to the edges of the prism P. A lens L is at the other end of the collimator tube C, whose length is adjusted until the slit is at the focus of the lens. The light from S is consequently converted into a parallel beam before it strikes the prism.

On emerging from the prism, each of the parallel colored beams

* For more detailed accounts, see C. E. K. Mees, *Photography* (New York, The Macmillan Company, 1937).

into which the original beam is resolved is brought to a focus at a different place in the focal plane BR of the eyepiece of the telescope T.

Suppose, for example, that the original light contained but two wave lengths, say, one blue and one red. The index of refraction of the prism for each of these colors is different; hence the prism resolves the original parallel beam into two parallel beams, one blue and one red. The length of the telescope T is adjusted until the objective lens O brings these parallel beams to foci in the plane BR, producing two **images of the slit,** one blue and one red.

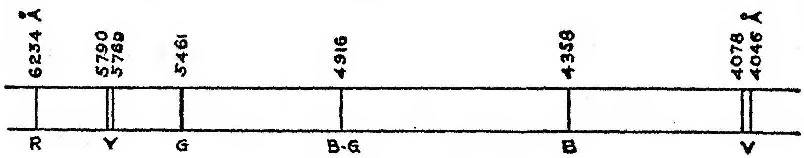


Fig. 587. Spectrum of Mercury

These are magnified by the eyepiece E and constitute the spectrum of the original light. The slit may be a centimeter or more long but it is always quite narrow, so that its images, whatever their colors, will always appear as lines.

Therefore, the spectrum produced by a spectroscope consists of a number of colored lines, **each line being an image of the slit formed by the light of a single color** (approximately). Figure 587 shows the prismatic spectrum of mercury.

The scale tube M contains at its end J a transparent scale and at the other end a converging lens. The scale is illuminated by a separate source of diffuse white light and its image, **reflected** from one face of the prism, is focused in the plane of the spectrum. The spectrum lines are therefore seen against the background of this fixed scale, on which their positions may be definitely located.

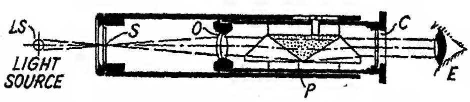


Fig. 588. Direct Vision Spectroscope

A **spectrometer** differs from a spectroscope in having the telescope free to turn about the axis of the prism, its position being read on a horizontal circle, and in having no scale tube. It may be used for various purposes besides the study of spectra, but it is particularly desirable when a grating is used to produce the spectrum instead of a prism.

A **spectrograph** is a spectroscope equipped with a plate or film holder so that photographs of the spectra may be taken.

In the **direct vision spectroscope** (Fig. 588), prisms of crown and flint glass are combined so that dispersion of the colors is obtained without deviation of the beam.

608. Kinds of spectra. Spectra are conveniently classified according to the following types:

I. *Emission spectra*, produced from light emitted by some self-luminous source with no intervening absorber. These are subdivided into:

(*a*) *Continuous spectra.* A spectrum is said to be continuous within a certain range of wave lengths when it contains all wave lengths within that range. Continuous spectra are produced by incandescent solids, liquids, and gases at high pressure. The incandescent tungsten or carbon of an electric lamp is the usual source of continuous spectra.

(*b*) *Bright line spectra*, emitted by atoms of elements in the gas state at low pressures and high temperatures. For example, if lithium chloride is placed in a Bunsen flame, it vaporizes and gives the flame a reddish hue; its visible spectrum consists of a bright red and a bright orange line.

(*c*) *Bright band, or fluted, spectra*, produced by the molecules of certain compounds when in the gas state at low pressure and high temperature, which disrupts the molecule. Under high resolution, the bands are found to be groups of fine lines closely crowded together on one side (the head) and less crowded on the other side. The head of a band may be on either the side toward the blue end of the spectrum or the side toward the red. The spectrum of CO shows both.

Ordinary bright line spectra are due to electron transitions within the atoms. Band spectra, on the other hand, are due to changes within the molecule. Thus a weak spark discharge between electrodes covered with a compound of mercury produces a band spectrum characteristic of the compound; but if the discharge is made sufficiently heavy to break down the compound, an ordinary bright line spectrum of its elements is obtained.

II. *Absorption spectra*, obtained by examination of light from a source yielding all wave lengths after this light has passed

through an absorbing medium. These are subdivided as follows:

(*a*) *Dark line spectra*, which appear as dark lines against a bright background. They are obtained on spectroscopic examination of light from a source yielding a continuous spectrum after that light has passed through an element in the gas state at a temperature somewhat lower than that of the source of light.

Kirchhoff's experiment demonstrates this: Set three Fisher burners in tandem and in line with the axis of the collimator. If we sprinkle common salt ($NaCl$) in the flames and examine the resulting yellow light, we find its visible spectrum to consist of the bright yellow line of sodium (really two lines very close together), and we note its position on the scale. Now let a carbon arc lamp be placed behind the bank of the Fisher burners so that the intense light from its horizontal + carbon shall pass through the three sodium flames and into the slit of the spectroscope. We now observe a continuous spectrum with a dark line exactly where the bright line of sodium was previously. If the sodium flames are removed, the dark line disappears, leaving the spectrum continuous. Evidently the sodium absorbs its own line from the continuous spectrum.

In general, **an incandescent gas absorbs from the light of a stronger source the colors (wave lengths) which the gas itself emits.** The phenomenon is an effect of resonance. The absorption lines are not totally black—only relatively dark as compared with the bright continuous background.

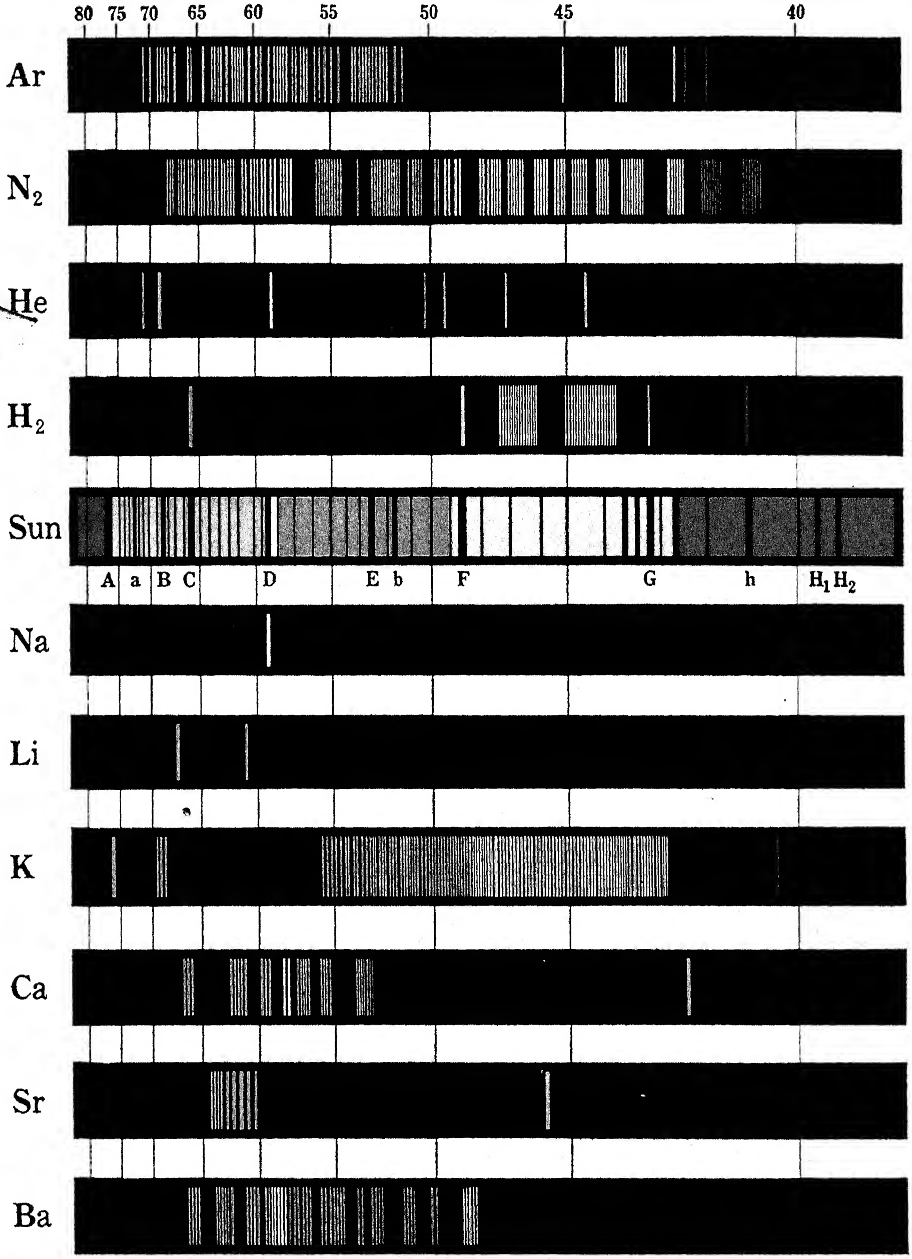
(*b*) *Dark band spectra* are of two kinds:

1. When light from a source emitting a continuous spectrum is passed through various substances in the gas state at correct temperatures and then examined spectroscopically, dark absorption bands are found in exactly the same positions as corresponding bright bands when the absorbing gas is the emitter (type I (*c*), above).

2. When light from a source emitting a continuous spectrum is passed through certain solids and liquids, the spectra of the emergent light shows broad dark regions which are not resolvable into fine lines and have edges that are not distinct.

Thus, if white light is passed through a dilute solution of blood, the spectrum of the emergent light shows charac-

PLATE OF SPECTRA



From *Inorganic Chemistry*, by H. P. Cady.
By permission of the author and McGraw-Hill Book Co., Publishers

teristic dark bands in the green-yellow region, against a background of the other colors.

609. Spectrum analysis. The various spectra have been found to be characteristic of the substances producing them, somewhat as fingerprints are characteristic of individuals. This fact of transcendent importance appears to have been discovered in 1854 by David Alter, a country physician of Elderton, Pa.*

Under different conditions of excitation, however, the spectrum of a given material will consist of different lines. Examination of a

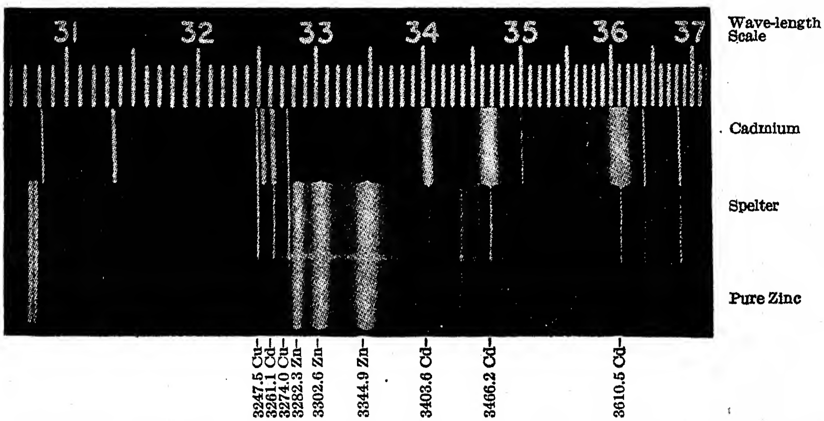


FIG. 589. Spectrogram Showing Cadmium and Zinc in Spelter. (Courtesy of Arthur H. Thomas Co.)

table of wave lengths will show the letters *a*, *s*, and *f* following certain wave lengths. These letters mean that the corresponding lines are obtained when the substance is excited in an electric **arc,** an electric **spark,** or a gas **flame,** respectively.

Amounts of a substance far smaller than can be recognized by ordinary chemical analysis can be detected by the spectroscope, or by long exposure with a spectrograph. Figure 589 is a spectrogram showing the lines of Cd and Zn in the spectrum of spelter.

While spectroscopic analysis was until recently entirely a qualitative process, it is now coming to be a very useful quantitative method. The amount of a substance is determined by the photographic density of the lines of the spectrogram, as measured

* *American Journal of Science and Art*, XVIII (1854), pp. 55–57.

with a densitometer. A "trace element" representing as little as 0.0001% of a mixture can be determined with a high degree of accuracy.

610. The Fraunhofer lines. In 1802, William Wollaston, a London physician, discovered in the spectrum of sunlight several dark lines, parallel to the slit. These lines are named, however, after Joseph von Fraunhofer of Munich, who discovered them independently in 1814 and whose careful study associated them with their corresponding bright line spectra. But it was G. R. Kirchhoff who explained them (1859) as **absorption lines of the elements in the atmosphere of the sun.**

The main body of the sun is so hot and dense that the light from it would yield a continuous spectrum. But the body of the sun is surrounded by an atmosphere of various elements in the state of

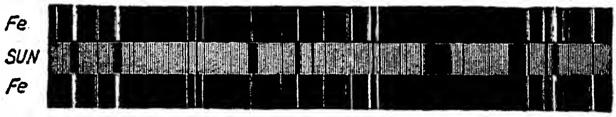


Fig. 590. Spectrum Showing Iron in Sun. (Courtesy of Allyn and Bacon)

incandescent gases. The light from the main body of the sun has to come through these gases, which then absorb their own lines from the continuous spectrum, leaving the dark, absorption lines in their places. Such absorption lines are just as characteristic of a substance as are the bright lines, since they occupy the same positions in the spectrum and hence give the same wave lengths as are given by the bright lines. Figure 590 shows some of the Fraunhofer lines (dark) and the corresponding lines due to iron (bright).

Besides being demonstrable in the laboratory (Sec. 608), this theory is beautifully confirmed at the time of a total eclipse of the sun. Just before totality occurs, the dark Fraunhofer lines are still distinctly visible; but at the instant that the moon completely covers the main body of the sun, the bright lines of the elements in its atmosphere **flash** out in exactly the positions of the corresponding Fraunhofer lines.

Fraunhofer gave letters to some of the stronger lines (see Fig. 593); but more than 20,000 of these dark lines have been recorded. In this way, 64 of the elements found on the earth are

found also in the *atmosphere* of the sun.* Helium was discovered in the atmosphere of the sun by Lockyer and Jansen 25 years before it was discovered on the earth.

Since spectroscopic analysis is made by means of the light from the substance, we can make the analysis wherever we can see the light. The distance makes no difference.

611. The Doppler effect. Before the advent of the spectroscope, astronomers could determine only the proper motion of stars, i.e., their motion at right angles to the line of sight. The spectroscope enables them to determine the other component—the motion along the line of sight.

This is done by observing the *Doppler effect:* If a star is moving toward the earth, the frequencies of its spectrum lines are increased, and they are shifted toward the violet end of the spectrum with

FIG. 591. Spectrum of Procyon Showing Doppler Effect. (Courtesy of Dr. V. M. Slipher)

respect to the lines of the same elements obtained in the laboratory. If the star is moving away from the earth, just the reverse is true, and the lines are shifted toward the red. The amount of the shift is proportional to the speed of the motion along the line of sight.

Thus Fig. 591 shows the Doppler effect in the spectrum of the star Procyon. The magnitude of the shift toward the violet indicates a speed toward the earth of about 18 mi/sec.

612. The complete electromagnetic spectrum. The wave lengths that produce the sensation of light are a relatively small part of the complete electromagnetic spectrum.

In 1800, Sir William Herschel found that when the bulb of a thermometer is held in a spectrum, its temperature rises more

* These are: H, He, Li, Be, B, C, N, O, F, Na, Mg, Al, Si, P, S, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, Rb, Sr, Y, Zr, Cb, Mo, Ru, Rh, Pd, Ag, Cd, In, Sn, Sb, Ba, La, Ce, Pr, Nd, Sa, Eu, Gd, Tb, Dy, Er, Tu, Yb, Lu, Hf, Ta, W, Os, Ir, Pt, Pb.—Charlotte M. Sitterly, *Proceedings of the American Philosophical Society*, LXXXI, No. 2 (1939).

rapidly as it is moved from the violet to the red end; and that beyond the red end it rises yet more rapidly than anywhere within the visible region. This radiation, whose wave lengths are *too long* for it to be visible, is called **infrared radiation,** or, less correctly, "infrared light." (By its definition, light properly refers to visible radiation only.)

The following year, J. W. Ritter discovered that silver chloride, which darkens when exposed to visible light, darkens still more rapidly if placed just beyond the violet end of the spectrum. This radiation beyond the violet, whose wave lengths are *too short* ~~for~~ it to be visible, is called *ultraviolet radiation*, or "ultraviolet light."

The ultraviolet and infrared regions of the spectrum cannot be investigated with an ordinary spectrometer, because the glass

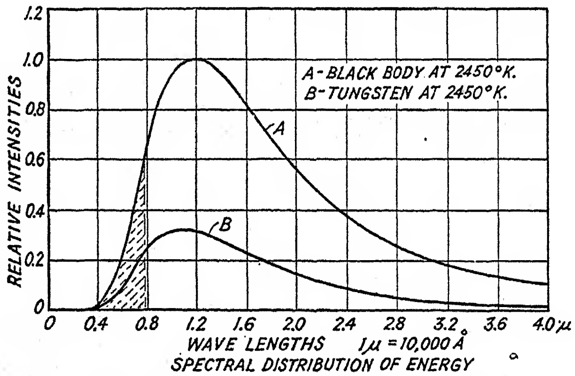


FIG. 592. Spectral Distribution Curves

lenses and prism absorb these wave lengths strongly. For the ultraviolet, a prism and lenses of quartz are generally used; but an infrared spectrometer uses a prism of rock salt, and concave mirrors instead of lenses.

A **concave grating,** i.e., a grating ruled on the surface of a concave mirror, may be used for the ultraviolet, visible, and infrared regions, all three, since it requires no lenses and its selective absorption is small.

Ultraviolet radiation produces chemical and biological effects, and its spectra are therefore usually photographed. Infrared radiation produces chiefly heating effects and is usually detected by means of a thermopile.

Figure 592 shows the distribution of energy radiated from a

black body at 2500°K. Abscissas are wave lengths, and ordinates are energy per cm^2 per sec per unit of $d\lambda$. It will be seen that the visible radiation (dashed) is a very small part of the total energy radiated, which is represented by the area under the curve.

But electromagnetic waves may be obtained in many other ways than from a hot body. By mathematical reasoning, James Clerk Maxwell concluded that an oscillating electric current should produce electromagnetic waves; and Heinrich Hertz, after Maxwell's death, succeeded in producing them.

X-rays, discovered by Roentgen in 1895, and the gamma (γ) rays given out by radioactive substances (Becquerel, 1896) are likewise known to be electromagnetic waves of extremely short wave lengths.

The complete electromagnetic spectrum, therefore, is much more extensive than the visible. In the language of music, it comprises about **50 octaves,** whereas the visible region (3900 Å to 7800 Å) embraces but **one octave.** The table below gives a general idea of the extent of the complete electromagnetic spectrum.

Complete Electromagnetic Spectrum

Kind of Wave	Order of Magnitude	Source
γ-rays	0.005 to 1.4 Å	Radioactivity
X-rays	0.05 to 100. Å	Impact of cathode rays
Ultraviolet	100. to 3900. Å	Ionized gases and hot bodies
Visible	3900. to 7800. Å	Ionized gases and hot bodies
Infrared	7800. to 3×10^6 Å	Hot bodies
Radio	0.1 to 10^6 cm	Oscillating electric currents

613. Spectra from prisms and from gratings. Spectra from prisms are called **irrational** because the position in the spectrum occupied by a line of given wave length depends upon the material of which the prism is made, and there is no rational mathematical relation between deviation and wave length (see Fig. 531). Hence wave lengths cannot be determined directly from prism spectra.

Spectra from gratings, on the other hand, do *not* depend upon the material from which the grating is made and have the simple relation of Eq. (428),

$$a \sin \theta = n\lambda \tag{445}$$

between wave length λ and deviation θ.

Hence on a grating spectrum the distance between lines is approximately proportional to the difference of wave lengths so that the spectrum is approximately normal. A **normal spectrum is one in which the angle of deviation is proportional to the wave length.**

In prismatic spectra, the short wave lengths are deviated most; but in grating spectra, the long wave lengths are. Hence the colors occur in reverse order in the two kinds of spectra.

A prism throws all the energy into a single spectrum, whereas

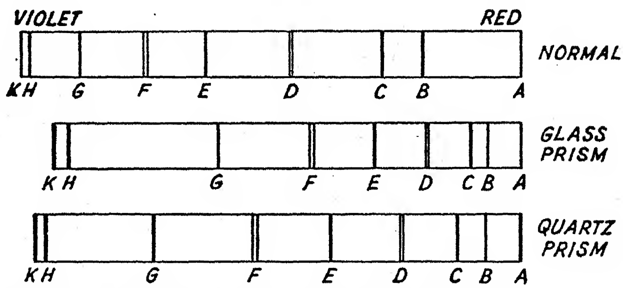


Fig. 593. Relative Positions of Lines in Spectra from Prisms and Gratings

a grating divides it among the various orders. In general, therefore, prismatic spectra are more intense (stronger) than are grating spectra.

Figure 593 shows the positions of the principal Fraunhofer lines when the spectrum of sunlight is made with a grating, a glass prism, and a quartz prism, respectively.

614. X-ray spectra. The nature of x-rays remained in doubt for a number of years. They did not appear to be reflected or refracted by ordinary mirrors and prisms. Assuming that this was on account of their extraordinarily short wave length, Max von Laue suggested in 1912 that if this were true, a crystal might serve as a diffraction grating for these waves, since presumably the atoms in a crystal are arranged in regular rows and layers which might have the proper spacing to enable them to play the role of the lines on an ordinary grating.

The experiment was immediately made by Frederick and Knipping, and it fully confirmed the suspicion that x-rays are very short electromagnetic waves. They passed a small beam of

x-rays perpendicularly into a crystal C (Fig. 594), behind which was placed a photographic plate P. Upon development the plate showed a diffraction pattern such as that of Fig. 595.

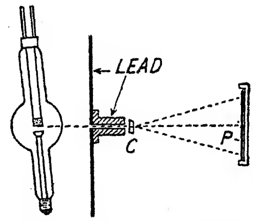


Fig. 594. X-Ray Tube Arranged for Making Laue Patterns

The central bright figure is where the plate was badly overexposed on account of the undiffracted beam. The surrounding bright spots ("Laue spots") are due to diffraction, and from their relative positions, the arrangement of the atoms in the crystal may be inferred.

The fundamental equation for a crystal used as a grating was derived by Sir William H. Bragg and his son, W. L. Bragg, of England, who also developed the method. It may be deduced as follows.

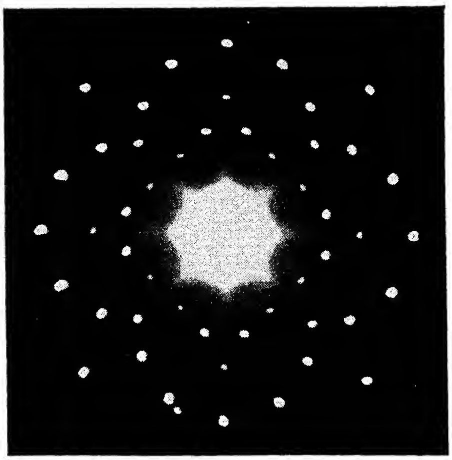


Fig. 595. Laue Pattern of Sodium Chloride

Let the small circles of Fig. 596 represent one arrangement of the atoms of a crystal in planes parallel to a cleavage face. XY is an incident, parallel beam of monochromatic x-rays making an angle θ with the atomic planes; and d is the distance between the planes, called the "grating space," or "lattice constant."

The index of refraction of matter for x-rays being practically unity, x-rays pass into it without sensible deviation. The energy of the incident beam enters the atoms of the crystal and is re-emitted from them as tiny wavelets. It may be shown* that these

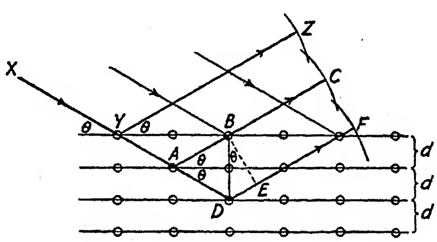


Fig. 596

* Richtmyer, *Introduction to Modern Physics* (New York, McGraw-Hill Book Company, 1928), p. 474.

wavelets have a random distribution, except in one direction in which they combine to form a new wave front ZF; and that the direction of this emergent beam is the same as if the incident beam had been reflected at the cleavage surface. The energy "reflected" at a single surface is trivial; but x-rays penetrate through thousands of atom layers, and the combined effect of the beams from all these reflections gives a resultant beam that is easily measurable.

If the waves reflected at A and at D are to yield a new wave front in the direction ABC, they must meet in phase; i.e., the path length ABC must differ from the path length ADF by a whole number of wave lengths.

The triangle ABD is isosceles, and therefore,

$$AB = AD.$$

Hence the path difference between ABC and ADF is DE, and DE must therefore be equal to a whole number of wave lengths, i.e.,

$$DE = n\lambda$$

where n is any integer.

But from the figure,

$$DE = 2d \sin \theta.$$

Therefore

$$2d \sin \theta = n\lambda \tag{446}$$

which is **Bragg's equation** for the crystal grating.

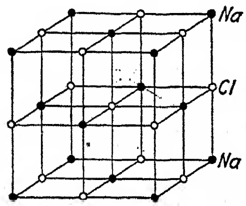


FIG. 597. Arrangement of Atoms in Sodium Chloride

By referring to Eq. (428), Eq. (446) is seen to be the same as the law of the diffraction grating, having the factor $2d$ in place of the grating space a. Wave lengths may be determined from this formula if d can be found for any one crystal. This is easily done for sodium chloride (Fig. 597).

A gram-molecule of NaCl has a mass of 58.454 gm and contains 6.064×10^{23} molecules (Avogadro's number). There are therefore

$$\frac{6.064 \times 10^{23}}{58.454} = 1.037 \times 10^{22} \text{ molecules/gm}.$$

The density of NaCl is 2.163 gm/cm³, and therefore there are

$$2.163 \times 1.037 \times 10^{22} = 2.243 \times 10^{22} \text{ molecules/cm}^3.$$

Since a molecule consists of two atoms, there are

$$4.486 \times 10^{22} \text{ atoms/cm}^3.$$

If these are arranged at the corners of elementary cubes, as the cubical form of NaCl crystals suggests, the number of atoms along *each edge* of the cubic centimeter would be:

$$\sqrt[3]{4.486 \times 10^{22}} = 3.553 \times 10^7 \text{ atoms.}$$

Hence the distance between atomic planes is:

$$d = \frac{1 \text{ cm}}{3.553 \times 10^7} = 2.814 \times 10^{-8} \text{ cm.}$$

Thus Eq. (446) enables us to evaluate the wave lengths of x-rays.

The x-ray spectrograph of the Bragg's is shown in principle in Fig. 598. X-rays from the target T are reduced to a very narrow beam by two slits S_1 and S_2 in plates of lead, and impinge upon the crystal grating C, from which they are reflected to the photographic film F. The table M carrying the crystal is rotated slowly about its axis perpendicular to the paper; and a line of the spectrum is recorded on the film F whenever θ has such a value that Eq. (446) is satisfied for one of the wave lengths λ radiated from the target.

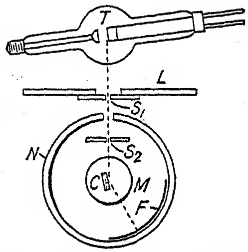


FIG. 598. Bragg's X-Ray Spectrograph

The spectrum is said to be of the first, second, third, etc., order according as n has the value 1, 2, 3, etc., respectively.

X-ray spectra are in general simpler than ordinary optical spectra, consisting usually of a continuous portion with a few prominent lines which are characteristic of the material of which the target of the x-ray tube is made. These characteristic lines occur in groups, or series, which are similar in all substances, and which have been given the names K, L, M, N, etc., series.

Figure 599 shows the K-series in the spectra of ten elements placed one above another. It will be noted that for each element the series consists of apparently two lines α, β, although actually

each of these has two components, and that the wave lengths get progressively shorter (and frequency, greater) as the atomic numbers increase, as is shown by the closer proximity of the lines to the undiffracted image *00*.

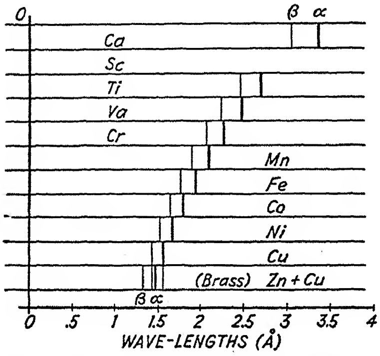


FIG. 599. Moseley's X-Ray Spectrograms (Redrawn)

615. Moseley's law. The similarity of the x-ray spectra of all the elements and their systematic change of wave length from element to element was discovered in 1913 by H. G. J. Moseley, some of whose spectrograms of the K-series are sketched in Fig. 599.

On plotting the square roots of the frequencies of these lines against the atomic numbers of the corresponding elements, the curve obtained (Fig. 600) was practically a straight line. The data for this purpose are given in the following table.*

Element	Atomic Number (Z)	Wave Length of K_α	$c\sqrt{\nu}$
Ca	20	3.36 Å	19.00
Sc	21	——	——
Ti	22	2.76	20.99
Va	23	2.52	21.96
Cr	24	2.30	22.98
Mn	25	2.11	23.99
Fe	26	1.95	24.99
Co	27	1.79	26.00
Ni	28	1.66	27.04
Cu	29	1.55	28.01
Zn	30	1.45	29.01

The numbers in the last column are the square roots of the frequencies of K_α multiplied by a constant factor as shown at the top of the column.

The curve actually curves upward slightly, but to a high degree of precision it is represented by the following equation: †

* From Siegbahn, *Spectroscopy of X-Rays* (New York, Oxford University Press, 1925), p. 87.

† Clark, *Applied X-Rays* (New York, McGraw-Hill Book Co., 1940), p. 110.

$$\sqrt{\frac{\nu}{R}} = \sqrt{\frac{3}{4}}(Z - 1) \tag{447}$$

where ν is the frequency;
R is the Rydberg constant; and
Z is the atomic number, i.e., the number of the place occupied by the element in the periodic table, counting from hydrogen as number one.

Stating this relation in words we have **Moseley's law:** The square roots of the frequencies of the lines of x-ray spectra have a simple linear relation to the atomic numbers of the corresponding elements.

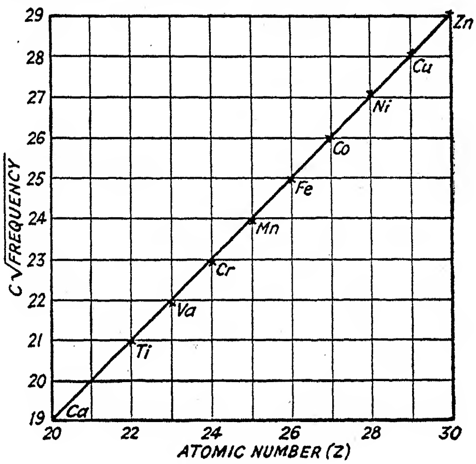


Fig. 600

This law showed that the important fact about an element is not its atomic weight but its atomic number, i.e., the ordinal number of the place it occupies in Mendeleeff's periodic table; and that nickel and cobalt should interchange places. It predicted the existence of elements then unknown which have since been found, such as hafnium and illinium. But most important of all, it showed that the elements are built up in a systematic way, and pointed to the nuclear charge as the determining factor.

616. Theory of spectra. It was observed by Balmer as early as 1885 that spectral lines occur in groups, or series, and that the frequencies of the lines in the visible spectrum of hydrogen were represented by the relation known as *Balmer's formula:*

$$\nu = cR\left(\frac{1}{n^2} - \frac{1}{m^2}\right) \tag{448}$$

where ν is the frequency of the radiation;
c is the velocity of light;
R is the Rydberg constant;

n may take values 1, 2, 3, 4, . . .; and
m may take values 2, 3, 4, 5, ...

But this remained a purely empirical relation until 1913, when Bohr, in a most brilliant application of Planck's quantum theory, showed that the equation could be derived rationally on the basis of certain hypotheses which were not in accord with classic electrodynamic theory.

Using the Rutherford planetary model of an atom, he assumed that there are certain stable orbits in which electrons may revolve about the positively charged nucleus without radiating. Whenever an electron is removed from one of these orbits by any means, such as collision with another particle, another electron is supposed to fall into the vacant position from one of the larger orbits, with the consequent emission of radiation of a definite frequency, depending upon the orbits from which and into which the electron fell.

Since each $-$ electron is attracted by the $+$ nucleus, work must be done on the system to remove an electron from a smaller to a larger orbit. Consequently, when an electron is in one of the larger orbits, the potential energy of the system is greater than when it is in a smaller orbit; and the system is somewhat unstable since in general a system tends to adjust itself so that its potential energy is a minimum. Hence electrons tend to fall back into the smallest orbit, which is called the **normal orbit** and corresponds to the stable state of the unexcited atom.

If W_1 is the potential energy of the system when an electron is in the smallest orbit, and
W_2 is the potential energy of the system when an electron is in a larger orbit;

then, when the electron falls back from the larger to the smaller orbit, the theory is that there is radiated an amount of energy $h\nu$ such that

$$h\nu = W_2 - W_1 \tag{449}$$

where h is Planck's universal constant ($= 6.55 \times 10^{-27}$ erg-sec).

The quantity $h\nu$ is called a **quantum** of energy, or a **photon,** and it is hypothesized that energy is radiated only in packets consisting always of a whole quantum, never with any fractional parts.

Quanta, however, are of different sizes, depending upon the value of the frequency ν.

On this basis, Balmer's equation may be derived.* It accounts for all the lines of the hydrogen spectrum. Similar formulas represent the series for ionized helium and doubly ionized lithium. The spectra of atoms involving more than one electron are much more complex. In most cases the series are approximately represented by **Rydberg's** empirical **formula**:

$$\frac{\nu}{c} = L - \frac{R}{(m + a)^2} \tag{450}$$

where L is the limit of the series,
m is any integer,
a is a constant characteristic of the particular series, and
R is the Rydberg constant.

The original idea of circular orbits has suffered various modifications. It is still a most helpful device for giving definiteness to a first discussion, but it is no longer considered to give a true picture of atomic structure. It is now customary to refer to the electrons as being at definite **energy levels** rather than in orbits.

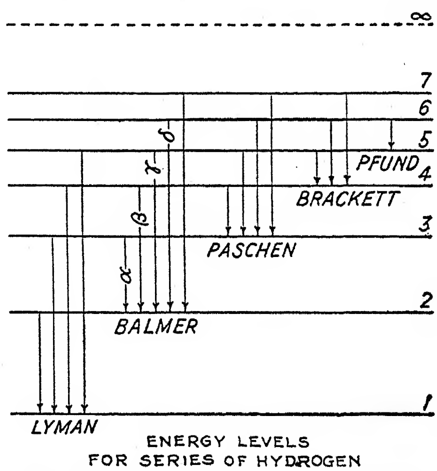


FIG. 601. Energy Levels for Hydrogen

Figure 601 shows diagrammatically how the theory accounts for the various series in the spectrum of hydrogen. When an electron falls from an energy level 3 to level 2 (i.e., when $n = 2$ and $m = 3$ in Balmer's equation), the line of lowest frequency H_α in the red ($\lambda = 6563$ Å) is radiated; if one falls from level 4 to level 2, H_β in the bluish-green ($\lambda = 4861$) is given out; and so on for the 31 lines that have been observed in the **Balmer series.**

Similarly, when $n = 1$ and m takes the values 2, 3, 4, . . ., we

* Foote and Mohler, *Origin of Spectra* (New York, Chemical Catalog Co., 1922), p. 16.

get the **Lyman series** in the ultraviolet; when $n = 3$ and m takes the values 4, 5, 6, . . ., we get the **Paschen series**; when $n = 4$ and m takes the values 5, 6, 7, we get the **Brackett series**; and when $n = 5$ and m takes the value 6, i.e., when an electron falls into energy level 5 from energy level 6, we get the single line which is the **Pfund series.** The last three of these series lie in the infra-red region of the hydrogen spectrum.

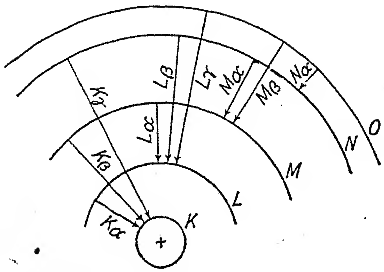


FIG. 602. Orbits of Inner Electrons

X-ray spectra are accounted for in the same way. When an electron is knocked out of the K-ring, or shell, and another falls into its place from the L-ring, the difference of the potential energies corresponding to these two configurations is radiated as the line K_α; if an electron falls from the M-ring into the K-ring, the line K_β is radiated; and so on, as shown in Fig. 602.

The frequencies of x-rays are so much greater than those of visible spectra that the electrons responsible for x-rays must lie very much closer to the nucleus in the normal, unexcited state of the atom than do those that produce optical spectra. X-rays are therefore believed to arise from inner electrons, whereas optical spectra arise from outer, or valency, electrons.

CHAPTER XLVI

OPTICAL INSTRUMENTS

617. The camera. The invention of the camera, under the name of "the camera obscura," is ascribed to the monk Roger Bacon (1214–92), who also invented gunpowder. In this early form the camera was used only for viewing and sketching. Photography was developed gradually later and is usually ascribed to Daguerre and Draper (around 1839).

The camera (Fig. 603) consists of a converging lens L in one end of a light-proof enclosure B, at the other end of which is a sheet of ground glass G. The enclosure is usually in the form of a bellows, so that the distance between L and G may be varied until the inverted image of an external object is focused on G. In the earlier forms, the image was sketched on the ground glass with a pencil. For photographic purposes, however, the ground glass is replaced, after focusing, by a sheet of transparent glass or celluloid, which is coated on the side next to the lens with a gelatine emulsion sensitized with bromide of silver.

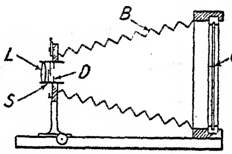


FIG. 603. Cross-Section of Camera

By opening the shutter S, the plate or film is "exposed" to the light that enters through the lens. This light causes a chemical change in the sensitive surface, which, on being properly developed and fixed, yields a permanent negative of the image produced by the lens.

The ordinary positive picture is made by placing the sensitized surface of a sheet of paper against the negative and permitting light to pass through the negative and act upon the paper. This print is then developed and fixed in a manner similar to that used in making the negative. Where the negative is less dense, more light will pass through than where it is more dense. Hence the lights and shadows of the negative are reversed on the print,

which is therefore a positive showing the highlights and shadows as in the original object.

The amount of light that reaches the sensitive plate depends upon the area of the diaphragm D and the time of exposure t. When the diaphragm, or stop, indicator is set at $f/16$, it means that the diameter of the opening is 1/16 the focal length. If set at $f/8$, the aperture is twice as large, and the area four times as large as before. The time of exposure in the latter case would be 1/4 that in the former. It is seen that the smaller the denominator, or f-number, the less the time required for a given picture. Hence the smallest f-number for which the lens will give a clear picture to the corners of the plate is called **the speed of the lens.**

In the "U.S. system," now seldom used, stop numbers were inversely proportional to the areas of the corresponding openings. U.S. stop 16 is the same as $f/16$.

The so-called "fixed-focus" camera is merely a camera having the lens set at the proper distance from the film to give a clear image of objects at the distance most commonly encountered, which is about 20 feet. For objects farther away or nearer than this, the image is not so clear. This is accomplished by using a single achromatic lens stopped down to about $f/11$, so that it is rather slow and produces good "snapshots" only in bright sunlight.

Rapid rectilinear lenses consist of two achromatic lenses having the iris diaphragm between them, their speed being usually about $f/8$, or twice that of the single achromatic type.

For a given distance of the object from the camera, **the longer the focus, the larger the image.** A **telephoto lens** is a long-focus lens designed to give a large image of a distant object. It is a combination of a converging lens with a diverging lens (Fig. 604). A **wide-angle lens** is one designed to give a large image of a near-by object. It is usually a large lens of short focus, stopped down to $f/16$ or less.*

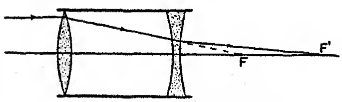


Fig. 604. Telephoto Lens

618. The eye. The eye is a marvelous camera with which most of Nature's creatures are equipped. Figure 605 shows a hori-

* Hardy and Perrin, *Principles of Optics* (New York, McGraw-Hill Book Company, 1932), p. 451.

zontal cross section of the right eye of a man; the various parts are labeled. It will be seen that the box of the camera is formed by the tough "sclerotic coat," which is opaque except in front of the crystalline lens, where it is transparent. This clear part is called the **cornea.** The space between the cornea and the lens is filled with a watery fluid, the **aqueous humor;** while the space behind the lens is filled with a transparent, jelly-like substance, the **vitreous humor.** The sclerotic wall is lined with the choroid coat, a dark layer of pigment cells and vascular tissue. The **optic nerve** enters at the rear and spreads out over the choroid, forming the **retina,** with its surface of microscopic rods and cones. The retina plays the role of the light-sensitive film in photography. At the center of the optic nerve where it enters the eyeball is the blind spot, and just to the right of this is the yellow spot.

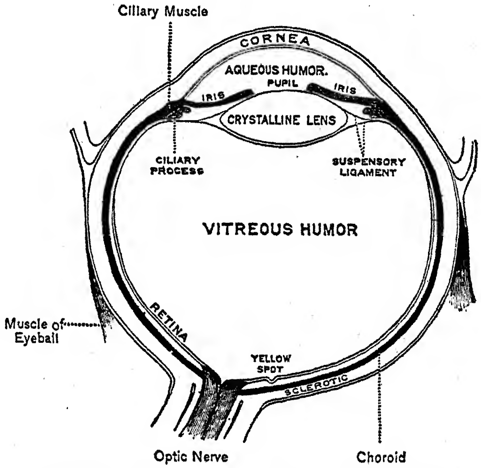


FIG. 605. The Human Eye. (Courtesy of D. C. Heath & Co.)

The image of an object is focused on the retina by varying the curvature of the crystalline lens. This is accomplished by the contraction or expansion of the ciliary muscle which surrounds the lens. The opening, or pupil, in front of the lens has its diameter changed automatically by the action of the iris, or colored portion of the eye. When the light is strong the iris contracts, making the pupil small, and vice versa.

The eye is most sensitive to greenish-yellow light ($\lambda = 5576$ Å), and only slightly sensitive to wave lengths in the extreme blue and red, as is shown by the visibility (or sensitivity) curve of Fig. 580.

619. Defects of the eye.

1. If the curvature of the cornea or of the crystalline lens is different in different planes, the spokes of a wheel will not all be

in focus at the same time, but will be slightly blurred in one section, as shown in Fig. 606. This defect is known as **astigmatism of the eye.** It is usually caused by the cornea and is measured by reflection of light from the surface of the cornea. Most people have a certain amount of astigmatism, which may not be at all serious. It is readily corrected by lenses having different curvatures in different planes to compensate for the differences in curvature of the eye system.

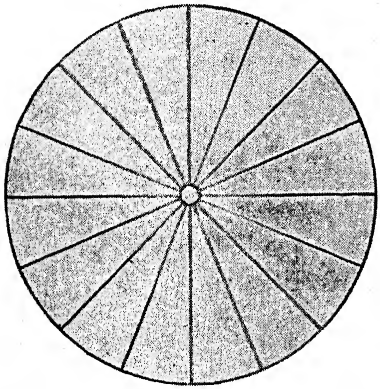


FIG. 606. Appearance of Wheel to Astigmatic Eye

2. If the focal length of the lens is too great for the length of the eyeball, only distant objects will be clearly focused on the retina. This defect is called farsightedness, or **hypermetropia.** It is corrected by converging lenses (Fig. 607).

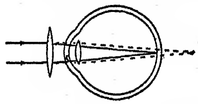


FIG. 607. Converging Lens Corrects Far-Sightedness

As one grows older, the **accommodation** of the lens, i.e., its ability to have its curvature changed, becomes less, so that at about 40 years of age one becomes farsighted and needs converging lenses. This defect is called **presbyopia.**

3. If the focal length of the lens is too short, only objects near by can be focused clearly. This defect is known as nearsightedness, or **myopia.** It is corrected by diverging lenses (Fig. 608).

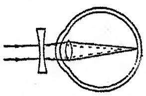


FIG. 608. Diverging Lens Corrects Near-Sightedness

620. The projection lantern. This apparatus (Fig. 609) is a camera reversed. A lantern slide *O*, which is a picture on glass or film, is intensely illuminated by a strong source of light *L*. Condensing lenses *C* serve to concentrate light upon the slide. An enlarged image *I* of the picture on the slide is projected onto a distant screen *S* by the **objective lens** *J* whose position may usually be varied by means of a rack and pinion.

621. Moving pictures. The projecting lantern for moving pictures is essentially the same as that described above for glass

slides. For this purpose, however, the positive pictures (usually 25 mm wide by 20 mm high) are made on celluloid film in strips about 1000 ft long, containing approximately 16,000 pictures.

Advantage is taken of the fact that **persistence of vision** is such that the visual impression on the retina lasts about 1/10 sec in average eyes. The motion picture projector is provided with a mechanism which presents a new picture before the condensing

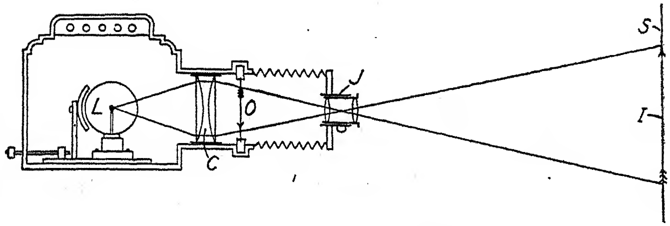


Fig. 609. Projection Lantern

lenses every 1/16 sec and holds it stationary for perhaps 1/32 sec, while a shutter uncovers the lens and permits the image to be projected on the screen.

Since 16 pictures are presented per sec and persistence of vision lasts 1/10 sec, the retinal impression of each picture remains until that of the next picture takes its place. The effect of smooth, continuous motion is thus produced.

When moving pictures are accompanied by sound, the sound record may be on separate phonograph disks whose motion is synchronized with that of the film. It has been found difficult to do this satisfactorily, however, and the best results are secured when the "sound track," or record, is made along the side of the picture film. For reasons of convenience in constructing the machine, the sound record for a given picture is about 14 in. below the picture to which it belongs.

The sound record consists of horizontal lines about 1/2 mm wide and 5 mm long, of varying opacity. A narrow beam of light is projected against this sound track and more or less passes through the film, depending upon the opacity of the record at that place. The light that gets through falls upon a photoelectric cell and thereby produces a varying current in an electric circuit. This current is amplified and actuates a loud speaker, which reproduces the original sound.

622. The magnifier. Any converging lens may be used as a magnifier.* It will be recalled from Case II of Sec. 558 that if the object is closer to the lens than the principal focus, the image is virtual, erect, and enlarged. If, therefore, we hold the eye close to a converging lens (Fig. 610) and bring the object AB up to a point between the principal focus F_2 and the lens, an enlarged image ab is seen.

For normal eyes, the distance of most distinct vision is about 25 cm (or 10 in.); hence the image is seen most clearly when it is about that distance from the lens. Eyestrain may often be relieved, however, without sensible change of magnification by focusing for the most comfortable working distance between final virtual image and eye, which is usually greater than 25 cm.

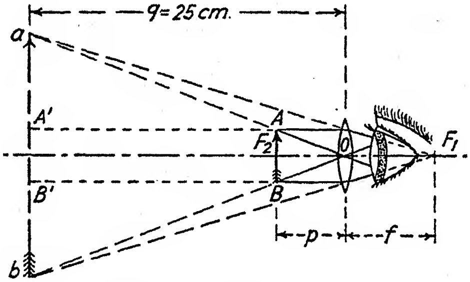


Fig. 610. Simple Magnifier

The magnifying power m of a lens is defined as **the ratio of the size of the image to the size of the object as seen without the lens,** each being at the distance of most distinct vision.

When the object is viewed without the lens, it is naturally placed at $A'B'$, in order to be most distinctly seen. Hence,

$$m = \frac{ab}{A'B'} = \frac{ab}{AB}.$$

But since $\triangle$ OAB and Oab are similar and AB is at the distance p from O,

$$m = \frac{ab}{AB} = \frac{25}{p}. \tag{a}$$

From the general lens formula, Eq. (422),

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

which, by the convention of signs (Sec. 558), becomes:

$$\frac{1}{p} - \frac{1}{25} = \frac{1}{f} \tag{b}$$

since $q = -25$.

* Sometimes called a simple microscope.

Solving Eqs. (a) and (b) simultaneously for m,

$$m = \frac{25}{f} + 1 \tag{451}$$

or, roughly,

$$m \doteq \frac{25}{f} \qquad \text{where } f \text{ is in cm.}$$

In words, **the magnifying power of a converging lens varies inversely as the focal length.** It was with such simple magnifiers of very short focal length, made by himself, that Anthony van Leeuwenhoek discovered bacteria in 1676.

622a. Resolving power. The resolving power of an instrument is its ability to show as separate two small adjacent elements in the structure of the object under observation. Quantitatively, the resolving power of a telescope is defined as the reciprocal of the angular separation of two points that are just resolved, i.e., shown to be two points instead of one. For a microscope the resolving power is defined as the reciprocal of the distance between two points that are just resolved.

FIG. 610a. Image of Point Source

On account of the wave nature of light, the image of a point source (star)* is not a bright point but a **diffraction pattern,** Fig. 610a. It is seen to consist of a bright central disk surrounded by alternate dark and bright rings resulting from the interference of wavelets coming from different points of the lens: the larger the lens, the smaller the diffraction pattern.

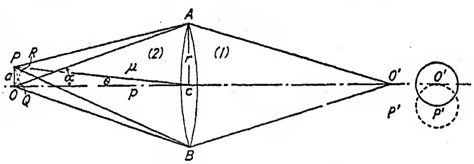


FIG. 610b. Resolving Power

As to when two points may be considered resolved, Lord Rayleigh has set up the arbitrary criterion that the center of one diffraction pattern shall fall on the first dark ring of the other. On the basis of this condition, we may deduce an expression for the resolving power of a lens as follows, Fig. 610b.

Let O be a point on the optic axis of the lens ACB and P a

point that is just resolved from O, OP being perpendicular to the axis. By the Rayleigh criterion, if O' is the center of the image of O and P' the center of the image of P, then P' will be on the first dark ring about O' and vice versa, as shown by the circles at the right of the figure.

If ${}_1\mu_2$ is the index of refraction of the medium to the left of the lens with respect to the medium to the right of the lens, ${}_1\mu_2 = \frac{v_1}{v_2}$ by Eq. 417. The time t for light to travel the distance PB in medium (2) will be $t = \frac{PB}{v_2}$. In the same time light would travel in the medium (1) a distance x such that

$$x = tv_1 = \frac{PB}{v_2}v_1 = \frac{v_1}{v_2}PB = {}_1\mu_2 PB.$$

Hence μPB is called the **equivalent path** in medium (1), usually air, and corresponds to the actual light path PB in the medium (2), μ being the index of refraction of the medium (2) with respect to medium (1), that is $\mu \equiv {}_1\mu_2$.

In order that O' shall be on the first dark ring about P', the equivalent light path by the route PBO' must be longer by one wave length (λ) than the path PAO', for then each wavelet from the lower half of the lens will arrive at O' one-half wave length behind the corresponding wavelet from the upper half of the lens so that they will nullify each other in pairs.

Stating this algebraically,

$$(\mu PB + BO') - (\mu PA + AO') = \lambda. \tag{a}$$

But

$$(\mu OB + BO') - (\mu OA + AO') = 0. \tag{b}$$

Subtracting (b) from (a),

$$\mu(PB - OB) + \mu(OA - PA) = \lambda. \tag{c}$$

Since OP is always small compared to OA or to OB,

$$PB - OB \doteq PR \qquad \text{and} \qquad OA - PA \doteq OQ. \tag{d}$$

From similar triangles, ORP and OCB,

$$\frac{PR}{OP} = \frac{CB}{OB} = \sin\alpha$$

$$PR = OP\sin\alpha. \tag{e}$$

From similar triangles, PQO and OCA,

$$\frac{OQ}{OP} = \frac{CA}{OA} = \sin \alpha$$

$$OQ = OP \sin \alpha. \tag{f}$$

Substituting from (d), (e), and (f) in (c),

$$\mu(PR) + \mu(OQ) = \lambda$$

$$2\mu OP \sin \alpha = \lambda$$

$$d \equiv OP = \frac{\lambda}{2\mu \sin \alpha}.$$

Sir George Airy (1834) showed that when the opening is circular the above value must be multiplied by the factor 1.22. Hence for circular lenses,

$$d = \frac{1.22\lambda}{2\mu \sin \alpha} = \frac{0.61}{\mu \sin \alpha}. \tag{451a}$$

In the case of a telescope, the medium on the object side of the lens is air so that $\mu = 1$; and the object distance OC ($\equiv p$) is great compared to the radius of the lens AC ($\equiv r$). Hence the angle θ subtended at the center of the lens by the distance OP between two points that are just resolved is

$$\theta \doteq \tan \theta = \frac{d}{p} = \frac{0.61\lambda}{p \sin \alpha}.$$

But
$$\sin \alpha \doteq \tan \alpha = \frac{r}{p}.$$

So that
$$\theta = \frac{0.61\lambda}{p\left(\frac{r}{p}\right)} = \frac{0.61\lambda}{r} = \frac{1.22\lambda}{D}$$

where D is the diameter of the lens.

Hence for a telescope, the resolving power $\equiv \dfrac{1}{\theta} = \dfrac{D}{1.22\lambda}$. (451b)

In the case of microscopes, the objects are usually not self-luminous and produce diffraction patterns themselves. Abbe has shown that under these circumstances the smallest object that can be distinguished has the diameter d such that

$$d = \frac{\lambda}{2\mu \sin \alpha}. \tag{451c}$$

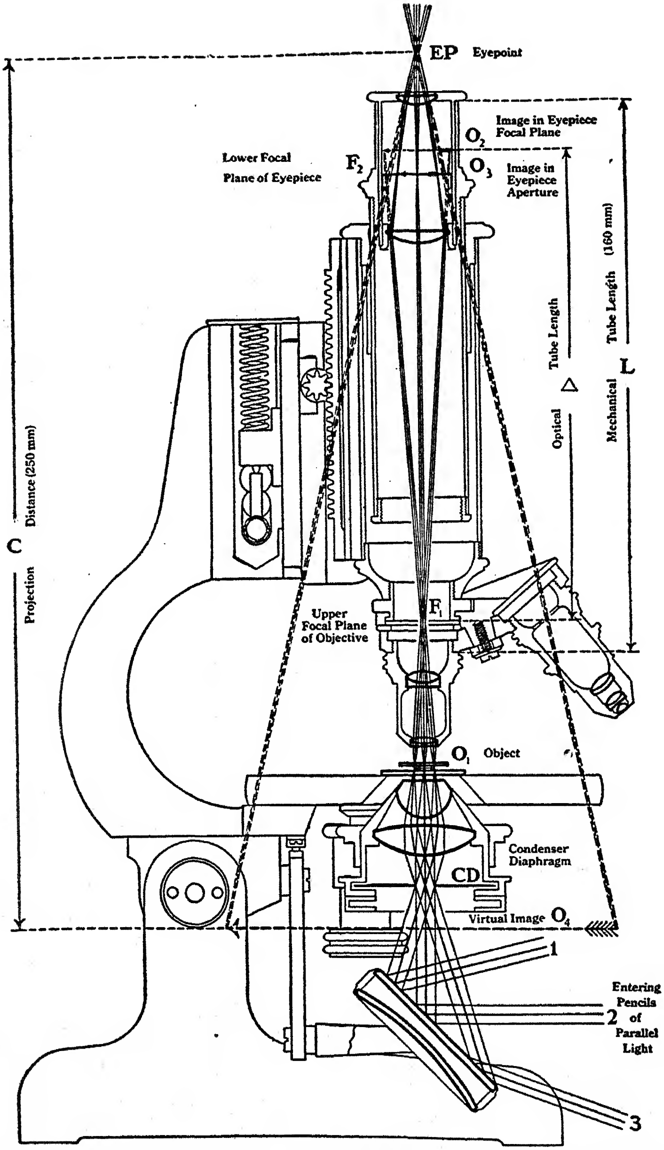


FIG. 611. The Miscroscope. (Courtesy Bausch & Lomb Optical Co.)

Hence

$$\text{the resolving power of a microscope} \equiv \frac{1}{d} = \frac{2\mu \sin \alpha}{\lambda}. \tag{451d}$$

623. The microscope. The principle of the microscope * is shown in Fig. 611. Its invention is attributed to Zacharias Joannides, about 1609.

The microscope will be seen to consist of two lens systems, each of which may be a single lens or several lenses. The one nearest the object is called the **objective,** and the one nearest the eye, the **ocular,** or eyepiece. The condenser and mirror are for illuminating transparent slides and are not parts of the microscope proper.

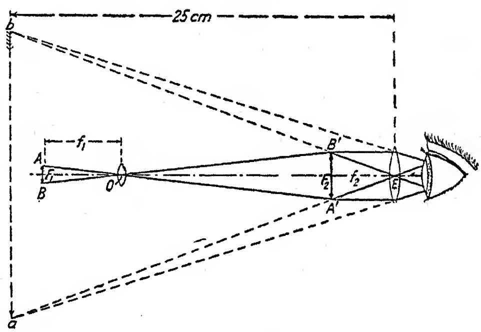


FIG. 612

Referring now to the simplified diagram of Fig. 612, the object AB is placed at a distance from O slightly greater than the focal length f_1 of the objective. Lens O produces a real image, inverted and enlarged, at $A'B'$, just beyond the principal focus F_2 of the eyepiece E, which, acting as the simple magnifier of the preceding paragraph, gives the virtual, enlarged image ab.

The distance from O to $A'B'$ is practically the tube length L of the microscope. Hence the magnifying power of the objective is approximately:

$$m_1 \doteq \frac{L}{f_1}.$$

The eyepiece magnifies the real image $A'B'$, its magnifying power m_2 being, by Eq. (451),

$$m_2 \doteq \frac{25}{f_2}.$$

Hence the total magnifying power of the compound microscope as a whole is:

* Sometimes called a compound microscope.

$$m = m_1 m_2 \doteq \frac{L}{f_1}\frac{25}{f_2} \doteq \frac{25L}{f_1 f_2} \tag{452}$$

where all measurements are in centimeters.

Manufacturers usually mark the magnifying power of each objective and eye-piece upon the piece itself. Thus an objective marked 80X means that the real image produced is 80 times as long as the object. Similarly an eye-piece marked 20X would yield a virtual image twenty times as long as the real image viewed by it. The total linear magnification of the object would therefore be $(80 \times 20) = 1600$ times, or 1600 diameters.

With visible light, magnifications as high as 3600X may be obtained, but 2500X is about the upper limit of useful magnification for most purposes. To secure magnifications higher than this, extremely short focus lenses must be used. Such focus is accompanied by small diameter, and small diameter of the objective means reduction of the illumination and of the resolving power of the microscope.

Resolving power is often quite as important as magnifying power, and is what limits useful magnification. The distinction between the two will readily be appreciated: **resolving power gives detail, magnifying power gives size.** Thus, we might make a lantern slide of Fig. 2 and project it on to a screen making it either two feet square or ten feet square. As far as our knowledge of the cathedral is concerned this would be useless magnification, for we could see in the smaller picture every detail of structure that would appear in the larger one. The detail was determined by the resolving power of the microscope with which the original photograph was taken; the projection lantern increases the size only.

With actual microscopes, the highest value of μ is 1.66 (monobrome naphthaline immersion) and the greatest value of α is 70°. The greatest practical numerical aperture is therefore

$$1.66 \sin 70° = 1.56.$$

Using this in Eq. 451c,

$$d = \frac{\lambda}{2\mu \sin \alpha} = \frac{\lambda}{2 \times 1.56} = \frac{\lambda}{3.12}.$$

Hence the smallest object that can be distinguished by direct

observation with a compound microscope cannot have a width less than about one-third the wave length of the light used.

Ultraviolet microscopes are now available achromatized for the 3650Å (ultraviolet) and the 5461Å (visible) lines of the spectrum of the mercury arc. Focusing may therefore be done by visible light and a photomicrograph made by the ultraviolet light, thereby greatly increasing the detail observable.

The limit of useful magnification has been enormously increased by the recently developed electron microscope.

623a. The electron microscope. As long ago as 1830, W. R. Hamilton pointed out that the path of a beam of light through a refractive medium in optics was analogous to the path of a particle through a potential field in mechanics.

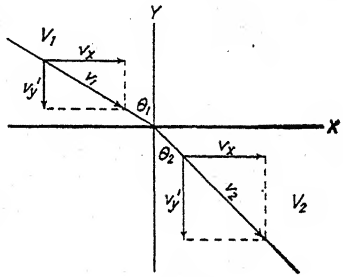


FIG. 612a.

In an elementary way this may be shown as follows: Suppose a beam of electrons, Fig. 612a, passes from a region of uniform electric potential V_1 into a region of uniform electric potential V_2, V_2 being greater than V_1. Let v_1 and v_2 be the velocities which an electron (charge e and mass m) would acquire in falling from a region of zero potential into the regions of potential V_1 and V_2, respectively. Then

$$\tfrac{1}{2}mv_1^2 = V_1e$$

$$\tfrac{1}{2}mv_2^2 = V_2e. \qquad \text{(a)}$$

As the electron passes into the second medium, the component v_x parallel to the plane of separation will remain unchanged, but the component perpendicular to that plane will change from $(v_1)_y$ to $(v_2)_y$.

We then have

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{\dfrac{v_x}{v_1}}{\dfrac{v_x}{v_2}} = \frac{v_2}{v_1}. \qquad \text{(b)}$$

But from (1)

$$\frac{v_2}{v_1} = \frac{\sqrt{V_2}}{\sqrt{V_1}} = \frac{\sin\theta_1}{\sin\theta_2}.$$

Therefore

$$\sqrt{V_1} \sin \theta_1 = \sqrt{V_2} \sin \theta_2. \qquad \text{(c)}$$

Comparing (c) with Eq. 418, it is seen that $\sqrt{V}$ is analogous to the absolute index of refraction of a medium.

These hypotheses were verified by H. Busch (1926) who showed that an axially symmetric magnetic field brings beams of electrons to a focus much as a lens brings beams of light to a focus. The same effect may be secured by an axially symmetric electric field. In the case of electric fields the paths lie in axial planes, but in magnetic fields the paths are spirals in consequence of Ampère's law. Such electric and magnetic fields are called **electron lenses** and their theory and utilization constitutes the subject of **electron optics.**

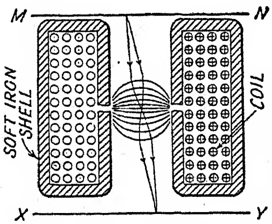


Fig. 612b. Magnetic Electron Lens

Figs. 612b and c show simple magnetic and electrostatic electron lenses, the fine curved lines representing equipotential surfaces. In each case, electrons from points of a pattern in the plane MN are brought to a focus in the plane XY which may be a fluorescent screen. The electrons may be emitted as photoelectrons from a semi-transparent surface at MN on which a picture is projected from above, or they may be supplied from a separate source and pass through a thin object mounted on a very thin supporting surface in the plane MN. Since air particles would deflect electrons, these lenses must be operated in a vacuum.

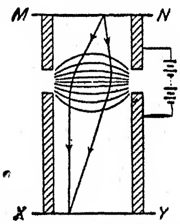


Fig. 612c. Electrostatic Electron Lens

Theoretically, for the electron microscope as for the light microscope, the least distance between two points that may be resolved is given by the relation $d = \lambda/(2\mu \sin \theta)$. In falling through the potential differences of 30,000 to 60,000 volts at which an electron microscope operates, the electrons acquire such velocities that their wave lengths as given by DeBroglie's relation (Eq. 472) are of the order of 1/10Å. On this account, the resolving power and useful magnifications of the electron microscope are enormously greater than those of the best light microscopes. In certain hydrocarbon molecules the atoms are about 2Å apart, so that it is quite within

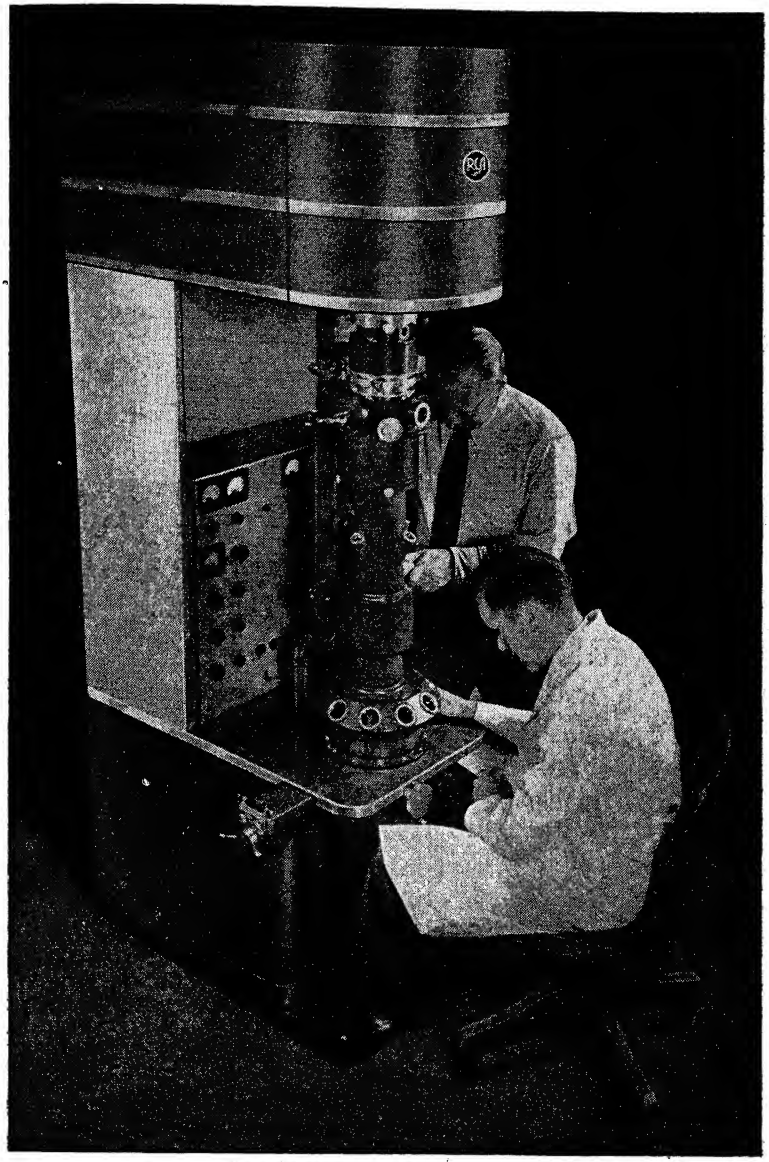


Fig. 612d. Electron Microscope.
(Courtesy RCA Electronic Research Laboratories)

the realm of future possibility to photograph individual atoms by means of the electron microscope.*

Fig. 612d shows the electron microscope developed by Dr. V. K. Zworykin and his coworkers of the Radio Corporation of America. A diagrammatic drawing of the instrument and of a similarly arranged light microscope is shown in Fig. 612e. In this instru-

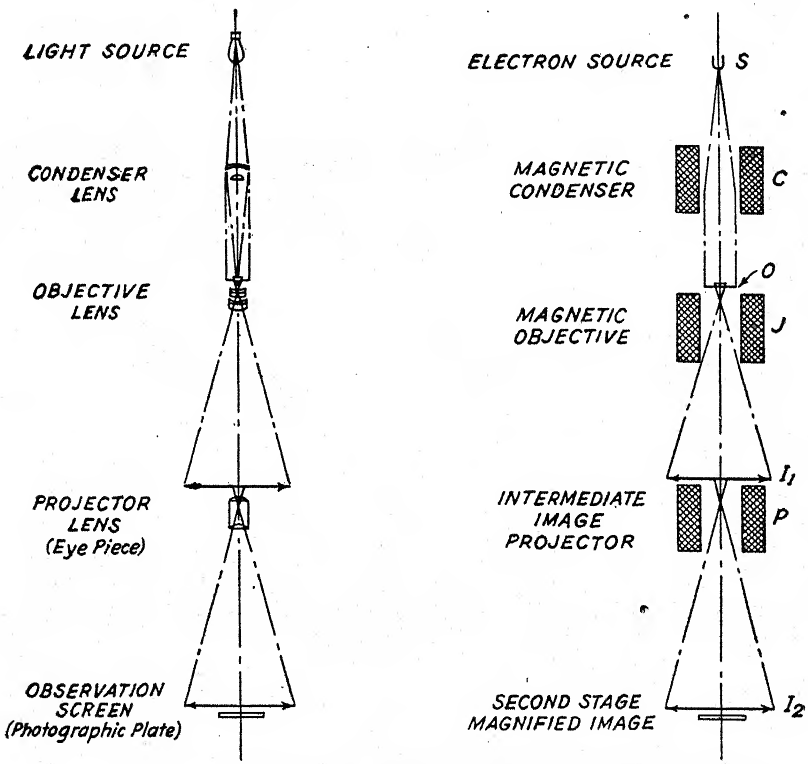


FIG. 612e. Comparison of Light and Electron Microscopes

ment magnetic electron lenses are used throughout because their focal lengths are more easily varied (by varying the current) than are the focal lengths of electrostatic lenses.

Electrons from a source S are concentrated by the condensing lens C upon the object which is mounted on a very thin supporting film at O. The objective J produces an intermediate image

* It is reported by Dr. Stuart Mudd that certain German scientists claim to have made visible the molecules of the organic compounds hemocyanin and edestin. (*Science News Letter*, October 26, 1940.)

at I_1, which may be viewed on a fluorescent screen. When this screen is removed, the projector coil P produces a further enlarged image on a second stage screen or photographic plate at I_2.

Direct magnifications of 30,000 diameters are secured as compared with the useful magnification of 2500 diameters with the light microscope; and so fine is the detail that photographic enlargements up to 100,000 diameters are advantageous.

624. The refracting telescope. Hans Lippershey, a spectacle-maker of Middleburg, Holland, is generally considered to have made the first telescope, in 1608. A telescope (Fig. 613) consists of an object glass, or **objective** O, and an **eyepiece** E. The function of the object glass is to collect light and produce a real image $A'B'$ of the object AB. The eyepiece serves as a simple magnifier and gives an enlarged virtual image ab.

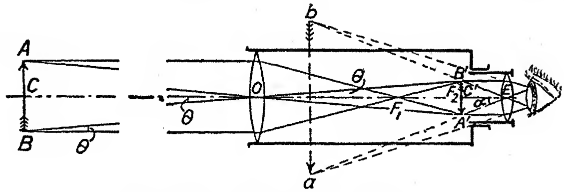


FIG. 613. Refracting Telescope

The eye estimates the size of an object by the angle which that object subtends at the eye. In this case it is more convenient to deal with half-angles.

If one should look at the object AB without the aid of a telescope, the half-angle which it would subtend at the eye would be sensibly the same as the half-angle θ which it subtends at the center of the objective, since the length of the telescope is negligible compared to the distance to the object. When viewed through the telescope, the half-angle subtended at the eye by the real image $A'B'$ is α. Hence the **magnifying power** of the telescope is:

$$m = \frac{\alpha}{\theta}.$$

Since θ and α are always very small angles, they are approximately equal to their tangents, so that

$$m = \frac{\tan \alpha}{\tan \theta}.$$

In an actual instrument, F_1 and F_2 are practically coincident (the figure has to be distorted here for clearness). Therefore,

$$m = \frac{\tan \alpha}{\tan \theta} = \frac{\dfrac{C'B'}{EF_2}}{\dfrac{C'B'}{OF_1}} = \frac{OF_1}{EF_2}.$$

But OF_1 is the focal length f_1 of the objective, and EF_2 is the focal length f_2 of the eyepiece. Hence the magnifying power is:

$$m = \frac{f_1}{f_2}. \tag{454}$$

Also, since F_1 and F_2 are practically the same point, the length L of a telescope is obviously given by the relation:

$$L = f_1 + f_2. \tag{455}$$

In telescopes used by engineers on levels and transits, an erect image is obtained by the introduction of an additional lens at the proper place between the objective and the eyepiece.

The resolving power of a telescope varies directly as the diameter of the objective; hence the constant endeavor to produce larger instruments. The largest refracting telescope today is the 40-in. refractor at the Yerkes Observatory, University of Chicago.

625. The reflecting telescope. Since the function of a telescope's objective is to collect light and bring it to a focus, a concave mirror may be used instead of a lens. The first reflecting telescope was made in 1668 by Sir Isaac Newton, who ground the mirror with his own hands. Its arrangement is shown in Fig. 614.

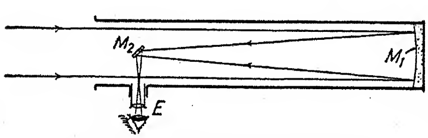


Fig. 614. Reflecting Telescope

The mirror objective is M_1, and a small plane mirror M_2 is placed in the path of the reflected rays so that the real image is viewed through the eyepiece E at one side.

The mirrors are generally made of glass of a low coefficient of expansion and are front-surface mirrors. This surface is coated usually with silver or aluminum. The former reflects about 81%

of the incident light, and the latter about 93%; the selective reflection of each is small. Consequently, a telescope having such a reflector may be used for the study of ultraviolet, visible, and infrared radiations, whereas one using a lens transmits little but the visible. Mirrors have two other advantages over lenses: only one surface has to be ground and polished, whereas a lens has two; and the mirror is more easily mounted, since a lens must be supported by its edge.

At present the largest telescope in the world is the 100-in. reflector of the Mount Wilson Observatory, but one 200 in. in diameter is under construction for the California Institute of Technology. Several years were required to grind and polish its mirror. When complete it will show probably 6,000,000,000 stars.

626. Parallax. If we look at a clock sidewise, it appears too fast or too slow—depending upon the side from which the observation is made. This apparent change of position of a body (in this case, of the clock hand) due to change of position of the observer is called **parallax.**

To avoid parallax in microscopes, telescopes, and other instruments, the cross-hairs are mounted on a ring, the **reticle,** and placed in the eyepiece tube so as to lie in the focal plane of the eyepiece, i.e., in the plane of the image $A'B'$ (Fig. 613). This eliminates any apparent motion of the cross-hairs with respect to the image, when the eye is moved across the eyepiece.

627. The Michelson interferometer. The beautiful colors seen on soap bubbles and thin oil films are due to the interference of light reflected at the two surfaces of such films. That this phenomenon could have any practical value might seem a remote possibility. But in 1887 Michelson utilized this very principle in his interferometer, which is still our most accurate device for the measurement of length, particularly the length of light waves. The instrument is shown in perspective in Fig. 615 and diagrammatically in Fig. 616.

The interferometer consists of two plane front-surface mirrors M_1 and M_2, and two plane parallel plates H and C. All should be plane to 1/20 of the wave length of sodium light, and are set perpendicular to the plane of the paper as shown.

H is half-silvered, i.e., the silver coating is so light that half

the incident beam is transmitted and the other half reflected. C is a clear glass plate cut from the same piece as H in order to have the same thickness. It is inserted in the path of the beam to M_2 to equalize the path length in glass, since the other beam traverses H three times.

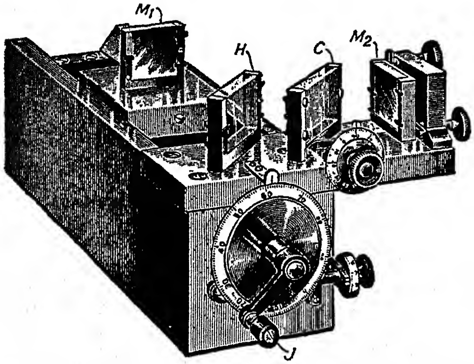


FIG. 615. Michelson's Interferometer. (Courtesy Franklin & Charles, Publishers)

Light from a steady monochromatic source S is made a parallel beam by the lens L and passes into the plate H. At the back surface, half is transmitted to M_2 and the other half is reflected to M_1, as shown by the full lines. M_1 and M_2 reflect these beams as shown by the dotted lines, and they enter the eye E.

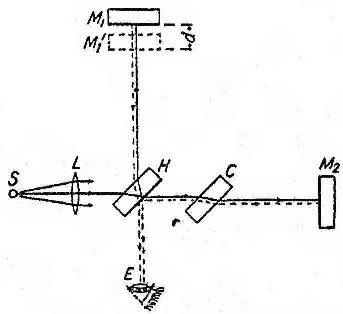


FIG. 616. Paths of Light in Interferometer

If the two path lengths are exactly equal, the two beams are 180° out of phase, because there is a phase change of that amount when waves in one medium are reflected against a denser medium,* and the field should be uniformly dark. But this is seldom the case, and usually the field is covered with light and dark interference bands, or "fringes." If M_1 is exactly parallel to the reflection of M_2, the bands will be circles. In general they are conic sections,† but by tilting M_2 they may be made practically straight. They are usually observed through a low-power telescope, and the field of view has the appearance of Fig. 617.

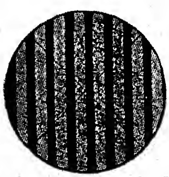


FIG. 617. Field of View of Interferometer

Mirror M_1 is mounted on a carriage that may be moved back and forth by a very accurate screw that is turned by the handle J.

If M_1 is moved a distance equal to $\lambda/4$, the length of the path of

* Edser, *Light for Students*, *op. cit.* p. 283. † Monk, *Light*, *op. cit.*, p. 146.

the corresponding beam is increased by twice this amount, or $\lambda/2$. Hence, where formerly there was a bright line there will now be a dark line, because the waves to this point will now be 180° out of phase and interference will take place. On moving M_1 another $\lambda/4$, the waves to this point will again be in phase, and the dark line will be replaced by a bright line because we again get reinforcement. So, as M_1 is moved very slowly, the lines appear to move across the field of view; and for each bright line that crosses the intersection of the cross-hairs we know that the mirror M_1 has been moved a distance $\lambda/2$.

Hence, if n bright lines pass the intersection of the cross-hairs, the distance d that M_1 has moved will be:

$$d = n\frac{\lambda}{2}. \tag{456}$$

One may count the passages of fringes across the intersection of the cross-hairs until n is a very large number, and the distance d may be read from the micrometer heads with great exactness. Consequently, **this gives our most accurate primary method of measuring the wave lengths of light.**

Using light of known wave length, such as the yellow line of sodium or the green line of mercury, the interferometer becomes a practical method of the greatest accuracy for other measurements of length. It was by this means that Michelson in 1894 determined the length of the prototype meter in terms of the wave lengths of the red, blue, and green lines of the spectrum of cadmium (Sec. 8). By means of a 20-ft interferometer attached to the 100-in. reflecting telescope at Mount Wilson, Michelson made the first measurements of star diameters. That of Betelgeuse was found to be 250 times and that of Antares 460 times the diameter of our sun.

628. The theory of relativity. The fact that **all motion is relative** was recognized in Newton's day, but he and all other physicists up to the time of Einstein assumed length, mass, and time to be independent and absolute quantities.

Consideration of the negative result of the Michelson-Morley experiment * led Einstein, in 1905, to the conviction that **true**

* G. F. Hull, *An Elementary Survey of Modern Physics* (New York, The Macmillan Company, 1936), p. 365.

length, mass, and time are meaningless terms—that we know only **measured** length, mass, and time. This inspired the enunciation of his two basic postulates:

1. The velocity of light, like that of sound and other waves, is independent of the motion of the source.

2. The velocity of light is the same to all observers, regardless of their own motion.

By mathematical reasoning from these, he developed a set of transformations called the **Lorentz transformations,** because they had previously been deduced by H. A. Lorentz for electromagnetic phenomena. Einstein assumed that when the algebraic expressions for all general physical laws were transformed from one set of moving rectangular axes to another set by these transformations, they retained their same algebraic form, provided the axes (or reference frames) were not accelerated with respect to one another.

This is known as the **special, or restricted, theory of relativity** and may be stated as follows: General physical laws have the same algebraic form relative to all Galilean frames of reference.*

The theory leads to the conclusion that if an observer had two identical meter-sticks and two identical clocks, and if one meter-stick and one clock were placed on a car and set in motion relative to the observer, the moving clock would appear to the stationary observer to run more slowly than the stationary clock. Likewise, if the length of the moving stick were parallel to the direction of motion, it would appear to the observer shorter than the stationary stick.

Thus, **time and length are not absolute values,** but depend upon the relative motion of the observer and the thing observed. Thus, a meter-stick on a car moving 161,000 mi/sec would appear only 50 cm long to an observer standing on the ground, but would appear 100 cm long to an observer riding on the car with it.

Einstein deduced also that a body whose mass when at rest is m_0 would at a velocity v have a transverse mass m such that

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{457}$$

where c is the velocity of light.

* A Galilean frame of reference consists of unaccelerated rectangular axes and Cartesian coordinates, as ordinarily used in analytical geometry.

This relation has been fully confirmed by the classic experiment of Bucherer,* in which he used electrons (β-rays from RaF) as the high-speed bodies. An electron whose velocity relative to a stationary observer is 0.9 that of light, would appear to that observer to have 2.29 times the mass m_0 of an electron at rest. Hence, **mass also is relative.**

But one of the most important conclusions of the Einstein theory is that **mass and energy are mutually convertible,** according to the relation:

$$E = mc^2. \tag{458}$$

This has been verified by the recent results of many workers on the transmutation of the elements (see Sec. 657). Thus, when $m = 1$ gm,

$$\begin{aligned} E &= 1 \times (3 \times 10^{10})^2 \\ &= 9 \times 10^{20} \text{ ergs.} \end{aligned}$$

That is,

$$1 \text{ gm mass} = 9 \times 10^{13} \text{ joules.}$$

Hence the laws of conservation of mass and of energy become one law.

In the Lorentz transformations, time enters as a fourth dimension (with a $\sqrt{-1}$ factor) on the same basis as length, breadth, and thickness. Minkowski has pointed out that we may be living in a four-dimensional "continuum," although we are able to visualize only the three dimensions of ordinary space.

Not satisfied with the limitation of unaccelerated axes, or frames of reference, Einstein pursued the study further and in 1915 brought out the **general theory of relativity:** All Gaussian coordinate † systems are essentially equivalent for the formulation of the general laws of nature.

The general theory leads to the fact that the geometry of a region is inextricably linked with the gravitational forces of the region, since the coefficients (g's) that determine the one likewise determine the other. From this standpoint it appears that gravi-

* Glazebrook, *Dictionary of Applied Physics* (London, Macmillan and Company, 1922), II, 350.

† Gaussian coordinates differ from Cartesian in that they may be families of curves instead of straight lines only. This is equivalent to permitting the moving bodies to have acceleration.

tational forces correspond to a curvature of the four-dimensional continuum, or region, in which we live. Light would accordingly pursue curved paths in gravitational fields, and such bending of light rays should be most noticeable in the vicinity of large masses of matter like our sun. This prediction has been confirmed by the apparent shift in the positions of stars photographed near the sun during several solar eclipses.

The general theory yields also a **law of gravitation,** more general than Newton's but also more complex. It consists of six differential equations which give Newton's law as a special case.

Einstein's law of gravitation accounts for the rotation of the orbit of the planet Mercury, which was not satisfactorily explained by Newton's law.

Still another test of the general law was suggested by Einstein. The theory predicts that the frequencies of atomic mechanisms considered as clocks should depend on the potential of the gravitational field in which they are situated. Spectral lines from the great stars should therefore be shifted toward the red (lower frequencies). Dr. W. S. Adams of the Mount Wilson Observatory has found this effect of the predicted magnitude in observations on the dark companion of Sirius.

At the present time, the theory of relativity of length, mass, and time is generally accepted. It represents one of the greatest advances in physical thinking since Newton, but it should not be regarded as a finality. Already modifications are being made by various writers and by Professor Einstein himself.*

PROBLEMS

1. If a movie projector shows 24 photographs per sec and a car wheel has 18 spokes, what are the three lowest speeds at which the wheel will appear to stand still? (This is known as the stroboscopic effect.)

2. Under certain light conditions the proper exposure for a picture was 0.1 sec when stop 8 was used. What would be the correct exposure using stop 32: (1) in the F-system; (2) in the U.S. system?

* References:

L. Bolton, *Introduction to Theory of Relativity* (New York, E. P. Dutton and Co., 1921).

A. Einstein (trans. by Lawson), *Relativity* (New York, Henry Holt and Co., 1921).

A. S. Eddington, *Space, Time, and Gravitation* (Cambridge University Press, 1935).

3. A telescope consists of an objective of 30 cm focal length and an eyepiece of 2 cm focal length. What is the length of the telescope (distance between lenses) when an object 180 cm from the objective is in sharp focus?

4. A microscope consists of an objective of 8 mm focal length and an eyepiece of 25 mm focal length. If the slide is 8.5 mm from the objective, what is the distance between the lenses and the magnifying power when the object is in sharp focus?

5. If the numerical aperture of an objective is 1.25, what is the distance between the two closest lines that can be distinguished by the sodium light? By visible light? By light of wave length, $\lambda = 2536\text{Å}$?

6. A microscope for photographing uses ultraviolet light of wave length 3600Å. What is the distance between the two closest lines that can be resolved by this microscope if its numerical aperture is 1.30?

CHAPTER XLVII

X-RAYS—RADIOACTIVITY—COSMIC RAYS

629. X-rays. The accidental discovery of x-rays by Wilhelm Konrad Roentgen in 1895 has revolutionized the work of the physician and the physicist. In 1894, Michelson had stated that the future of physics lay in the sixth decimal place—meaning

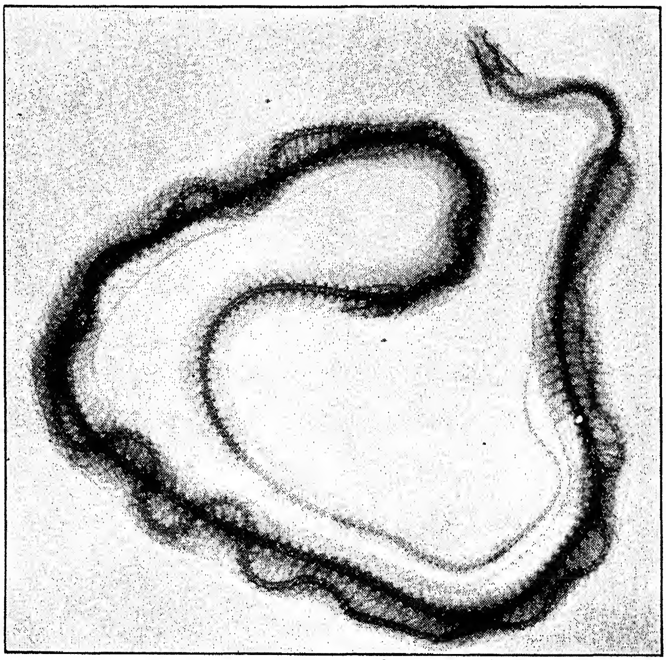


FIG. 618. Radiograph of King Snake That Has Swallowed Another Snake Longer than Himself. (Courtesy of Dr. Howard A. Kelly, Johns Hopkins University)

that physics was practically a completed science and that probably nothing remained but to increase the refinement of measurements. No surmise could have been farther from the fact. The discovery of x-rays opened unsuspected and still unlimited fields for new research and gave to physics a stimulus hardly equaled since the time of Newton.

X-rays have been shown (Sec. 614) to be electromagnetic waves similar to light but shorter. They have the following useful properties:

1. They proceed in straight lines and cast shadows.
2. They readily penetrate many materials such as cloth, wood, flesh, and the light metals such as aluminum.
3. They do not penetrate readily the heavy metals, such as lead and tungsten.
4. They act on photographic films as does light.
5. They cause zinc sulphide, platinum-barium-cyanide, and other substances to fluoresce.
6. They produce ionization of gases.

Figure 618 is a radiograph showing the "survival of the fittest": a king snake has swallowed another snake considerably longer than himself, as is shown by the loop made by the tail of the latter.

The ability of x-rays to penetrate matter and affect photographic films enables the surgeon to locate accurately foreign objects in the body, and the engineer to detect flaws and cavities in castings and welds. It permits definite determination of the presence and progress of tuberculosis and other diseases. Irradiation with x-rays has become a standard treatment for certain types of cancer and other abnormalities; and by it the process of evolution in certain forms of life has been speeded up enormously. To the physicist, the ability of x-rays to ionize a gas is perhaps their most important property. Their discovery marks the beginning of the "Modern Physics."

630. The electron-volt. In these days it is customary to express the energy of electrons, protons, quanta, etc., in terms of a unit called the *electron-volt.* The meaning of this term is easily understood.

If the potential difference between an anode and a cathode is V volts, then the work W required to move the quantity e of one electron from the anode to the cathode is:

$$W = Ve$$

and this is the **potential energy** that the electron has when it leaves the cathode.

As it falls toward the anode, this potential energy is gradually

transformed into kinetic; and if the electron reaches the anode without collision, its energy is all kinetic and has the value

$$\tfrac{1}{2}mv^2.$$

Hence,

$$W = Ve = \tfrac{1}{2}mv^2 \tag{459}$$

where m is the mass and v the velocity of the electron.

In words, **an electron-volt is the kinetic energy of an electron that has fallen freely through a potential difference of one volt.**

Since e means the quantity of electricity represented by 1 electron,* it follows that these energies in electron-volts will be numerically equal to the potential difference in volts through which the electron has fallen or could fall.

$$\begin{aligned} 1 \text{ electron-volt}^{\dagger} &= 1 \text{ volt} \times 1 \text{ electron} \\ &= \tfrac{1}{300} \text{ statvolt} \times 4.80 \times 10^{-10} \text{ statcoulomb} \\ &= 1.6 \times 10^{-12} \text{ erg.} \end{aligned}$$

631. Discovery of radioactivity. The fluorescence of the earlier forms of x-ray tubes led Henri Becquerel to think that perhaps fluorescent substances emitted x-rays. He therefore began an examination of such materials and found, in 1896, that the salts of uranium suspended above a photographic plate, in its plate-holder, produced a shadow photograph of opaque objects lying on the plate-holder, quite similar to an x-ray photograph.

To his surprise, however, he discovered that uranous salts gave the same results as uranic salts, although the former are not fluorescent. Clearly, therefore, the fluorescence was not an essential accompaniment of this newly discovered phenomenon.

Radioactivity is the property of spontaneously emitting rays capable of penetrating matter opaque to ordinary light. These new rays were called "Becquerel rays," and like x-rays they produced ionization of the surrounding air.

632. Radium and other radioactive elements. At the time of Becquerel's discovery there were in Paris two young French physicists, Pierre Curie and his wife, Marie Sklodowska Curie. They at once began a systematic examination of all sorts of substances

* Electron, as originally used by Stoney, meant quantity without regard to sign.

† The word "electron" is often omitted and energy stated as so many "volts," when electron-volts is meant.

for this property of radioactivity, using the electrometric test (i.e., the ability of a substance to discharge an electroscope by ionization of the air) as a measure of their radioactivity.

Of the then known elements, it was found by Mme. Curie * that only thorium possessed radioactivity comparable to that of uranium. Certain ores of these substances exhibited greater activity than the elements themselves. The Curies hypothesized that this was due to the occurrence in the ores of a small quantity of a substance more active than either element, and that this substance was thrown away in the refining process.

On this theory they made a most exhaustive analysis of **pitchblende,** the ore of uranium, and in 1898 discovered two new elements, **polonium** and **radium,** the radioactivity of the latter being a million times as great as that of uranium.† The former occurs with bismuth and the latter with barium, the separations being difficult—particularly that of polonium.

Radium is generally used in the form of the chloride or the bromide, but in 1910 Mme. Curie succeeded in preparing pure radium. It is a silvery-white metal similar to barium, and it tarnishes rapidly in the atmosphere. Its atomic weight is 225.97 and its melting point, 700°C. One ton of pitchblende contains about 0.2 gm of radium and 1/5000 this amount of polonium.

About 1900, Debierne, working in the laboratory of the Curies, discovered in the iron group and in rare earths another radioactive element, **actinium.** In 1907, N. R. Campbell found potassium and rubidium slightly active; ‡ and in 1932 Langer and Riatt observed radioactivity of beryllium.

Of the 12 or 15 elements essential to every living cell, potassium is the only one definitely radioactive. Rubidium and caesium, which are also slightly radioactive, may be substituted for potassium; but no non-radioactive element has been found that can take its place.§

Radioactivity is widely distributed throughout the waters and the crust of the earth. Some 40 elements have been found to

* And independently by Schmidt.

† Marie Curie, *Pierre Curie*, trans. Kellog (New York, The Macmillan Company, 1923), p. 114.

‡ Rutherford, *Radioactive Substances and Their Radiations* (New York, The Macmillan Company, 1913).

§ Burns, *Biophysics* (New York, The Macmillan Company, 1929), p. 161.

possess the property, and it is not improbable that at least feeble radioactivity is a property of all matter.

An artificial radioactivity usually of short duration, somewhat different from the spontaneous natural radioactivity described above, has been produced by many experimenters (see Sec. 640).

633. Alpha-, beta-, and gamma-rays. From a study of the penetrating power of Becquerel rays, Rutherford concluded that they consisted of three different types of component rays to which were given the names, α-rays, β-rays, and γ-rays.*

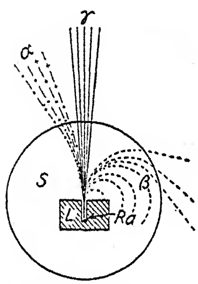


FIG. 619. Radiations from Radium

These rays are readily separated by an experiment attributed to Mme. Curie (Fig. 619). A small piece of radioactive material *Ra* is placed at the bottom of a hole in a block of lead *L*, from which the radiation escapes by a narrow slit. The block of lead is placed between the poles of an electromagnet *S*—only the south pole, behind the block, being shown in this figure. Before the current is turned on there is but a single beam, vertically upward, but after the magnet is energized this beam is split into three parts as shown.

As the magnetic flux is perpendicular to the paper and away from the observer, application of Fleming's motor rule shows that the deflections have the directions that we would expect if the α-particles have + charges, the β-particles − charges, and the γ-rays no charge at all. Examination with an electroscope proves this to be the case.

Rutherford secured similar separation by passing the discharge between parallel plates highly charged + and −.

All three of these radiations darken photographic plates and produce ionization of a gas. They emanate from the nucleus of the atom.

The *alpha-rays* have been shown by Rutherford and Royds to be the nuclei of helium atoms.† Accordingly they have an atomic mass of 4 and a positive charge of 2 protons. It is therefore better to speak of them as *particles* than as rays. Though

* Rutherford, *op. cit.*, p. 115.

† Rutherford, Chadwick and Ellis, *Radiations from Radioactive Substances* (New York, The Macmillan Company, 1930), p. 48.

expelled with velocities from 1.4×10^9 to 2.2×10^9 cm/sec, they penetrate only about 7.5 cm in a standard atmosphere and are completely stopped by a sheet of paper. However, they have great ionizing power because of their relatively great mass (7200 times that of a β-particle).

Beta-rays are likewise particles, for they are high-speed electrons, being expelled with velocities from 1/3 to 99.8% of the velocity of light, i.e., from 10^{10} to 3×10^{10} cm/sec. Those of the highest speeds have energy of the order of 8×10^6 electron-volts, whereas α-particles have 5 times this energy, which accounts for their greater ionizing power. The great penetrating power of β-particles is due to their high speed and small size.

Many substances fluoresce under bombardment by α- and β-particles, particularly zinc sulphide and some diamonds. Fluorescent paint is usually radiothorium mixed with zinc sulphide.

Gamma-rays are undeflected by an electric or a magnetic field; hence they carry no electric charge. They are electromagnetic waves of the same nature as light but of much shorter wave lengths. By crystal diffraction measurements such as are used for x-rays, the γ-rays of *Ra* have been found to consist of some 20-odd wave lengths from 28 to 1365 X.U.

$$1 \text{ X.U.} \equiv 1 \text{ x-ray unit} \equiv 0.001 \text{ Å} = 10^{-11} \text{ cm.}$$

Gamma-rays are in a sense a secondary effect, since they arise from the readjustment of the atomic structure after the expulsion of an α-particle or a β-particle from the nucleus. Gamma-rays have great penetrating power, being able to pass through as much as 20 cm of lead. Like x-rays, they produce ionization by causing the ejection of photoelectrons in accord with Einstein's photoelectric equation. They may be used for the same purposes as x-rays in the treatment of disease and the production of radiographs.

The relative ionizing powers of α-, β-, and γ-rays are approximately as 10,000 is to 100 is to 1.

Their relative penetrating powers, on the other hand, are roughly as 1 to 100 to 10,000, respectively.

634. Methods of detecting the radiations.

1. By scintillations. In 1902 Sir William Crookes devised the "spinthariscope" (Fig. 620) for detecting and counting α-particles.

It consists of a metal tube T about 1 in. in diameter and 2 in. high, at the bottom of which is a zinc sulphide screen S. A few millimeters above this is supported a piece of radioactive material R. Each α-particle shot out from R produces a bright flash, or scintillation, where it strikes the zinc sulphide screen. These scintillations are easily observed by means of a lens L at the top of the tube, when the apparatus is in a darkened room. By proper arrangements for limiting the fraction of the total radiation that reaches the screen, the scintillations may be readily counted.

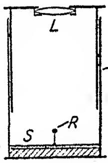


FIG. 620. Crooke's Spinthariscope

2. *By the Geiger counter.* This counter (Fig. 621) is an improvement on an earlier form devised by Rutherford and Geiger which counted only α-particles.

The radioactive material R is mounted at the center of a spherical container S in which is a small opening O by which the particles pass into the ionization chamber T through a thin mica window W, 2 mm in diameter. The ionization chamber consists of an outer metallic tube T connected to the + terminal of a high voltage battery C. Along the axis of T and well insulated from it is a sharp-pointed metal rod P connected to the grid of a triode, as shown.

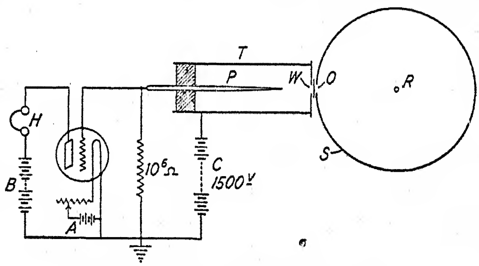


FIG. 621. The Geiger Counter

The voltage of C is adjusted until the gas in T is in a critical state just below the sparking potential. The ionization then produced by the entrance of either an α- or a β-particle is sufficient to cause at the point of the rod a small spark (visible with a microscope), which shows that the ionization has suddenly increased greatly. This permits electricity to flow much more readily from the casing to the point, thereby increasing momentarily the charge on the grid. This in turn increases the plate current and registers a click in the ear phones, for which a loud-speaker may be substituted.

If a string galvanometer is used instead of the phones, its deflections may be recorded on moving photographic film, as shown

in Fig. 622, where each spur represents the arrival of an α- or a β-particle. With this device as few as one particle per second may be counted.

3. By the Wilson cloud chamber. In 1912, C. T. R. Wilson observed that the ions produced in dust-free gas by α- and β-particles act as nuclei on which water condenses from a supersaturated atmosphere forming visible droplets, thus making visible the path of the ionizing particle. His apparatus, known as the *Wilson cloud chamber*, has been developed by Professor C. T. Knipp into the very convenient form shown in Fig. 623.

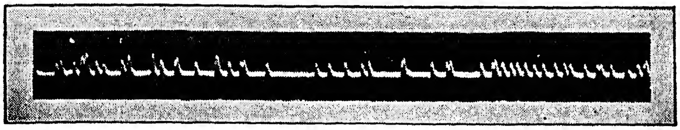


Fig. 622. Record from Geiger Counter

A flat-topped glass vessel G is closed at the bottom by a heavy rubber bulb B and nearly filled with water colored black with nigrosine. A small piece of a radium salt is enclosed in a thin-walled tube at R.

When the bulb B is squeezed, the air above the water is compressed. It should be held in this condition for a few seconds to allow the heat of compression to be dissipated. It is then released suddenly. The air, expanding adiabatically, is thereby cooled and consequently supersaturated. In this condition the vapor condenses in visible droplets upon the ions formed by the particles that are being continually shot out by the radioactive material. These droplets mark the paths of the α- and β-particles and of the ions due to γ-rays and x-rays. (The 110-volt d-c circuit is to clear the atmosphere of ions after each expansion.)

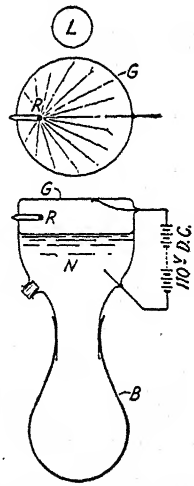
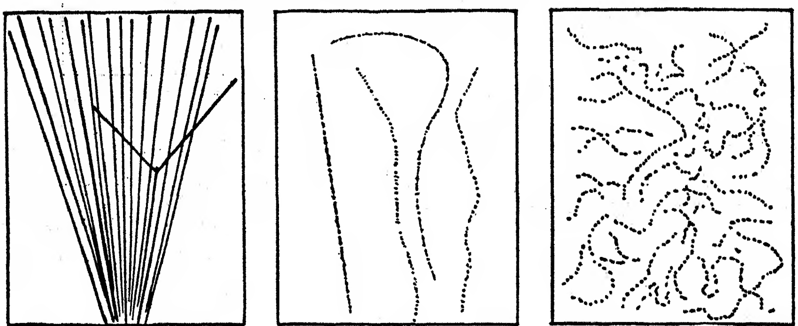


Fig. 623. Knipp's Cloud Chamber

These paths may be photographed by a brilliant spark; but on account of the very vigorous ionization produced by α-particles, their paths are made plainly visible by an ordinary incandescent lamp L behind G.

In Fig. 624 are shown sketches from photographs of the tracks made by α-particles, β-particles, and the ions produced by γ-rays. These tracks, or trails, consist of the droplets of water that have condensed on the ions formed by the three types of radiation. Those due to α-particles are broad on account of the great ionizing power of these rays. The tracks of β-particles are narrow and not continuous: high-speed β-particles give straight tracks, but



(α) (β) (γ)

FIG. 624. Cloud Chamber Tracks

the tracks of low-speed ones are sinuous. The ions (electrons) ejected from the atoms of the gas by γ-rays or x-rays wander indiscriminately, because of little energy and many collisions.

635. Radioactive transformations. The discovery by Rutherford in 1900 that thorium, and later that radium and actinium, emitted a radioactive "emanation," led Rutherford and Soddy in 1903 to advance the **transformation theory.** This theory states that radioactive elements are not stable structures but undergo spontaneous disintegration by the expulsion of α- and β-particles.

The subsequent proof that the α-particle is an ionized atom of helium (i.e., an atom of helium that has lost two electrons and has, therefore, an atomic mass of 4 and a charge of $+2e$), together with much other evidence, furnished the basis for the enunciation by Russell (and by Soddy and Fajans) of the *Displacement law:* The expulsion of an α-particle from the nucleus of an atom leaves a new atom whose atomic mass is less by 4 and whose atomic number * is less by 2 than those of the parent atom. Also, when an atom loses a β-particle, the atomic mass remains unchanged; †

* The atomic number equals the number of $+e$ charges in the nucleus (Sec. 498).
† Because the mass of the ejected β-particle is negligible.

but the atomic number increases by 1, since the loss of $-e$ in the nuclear charge is equivalent to an increase of $+e$.

In accordance with this law, the 40-odd radioactive elements are found to fall into three families, or series: the uranium-radium

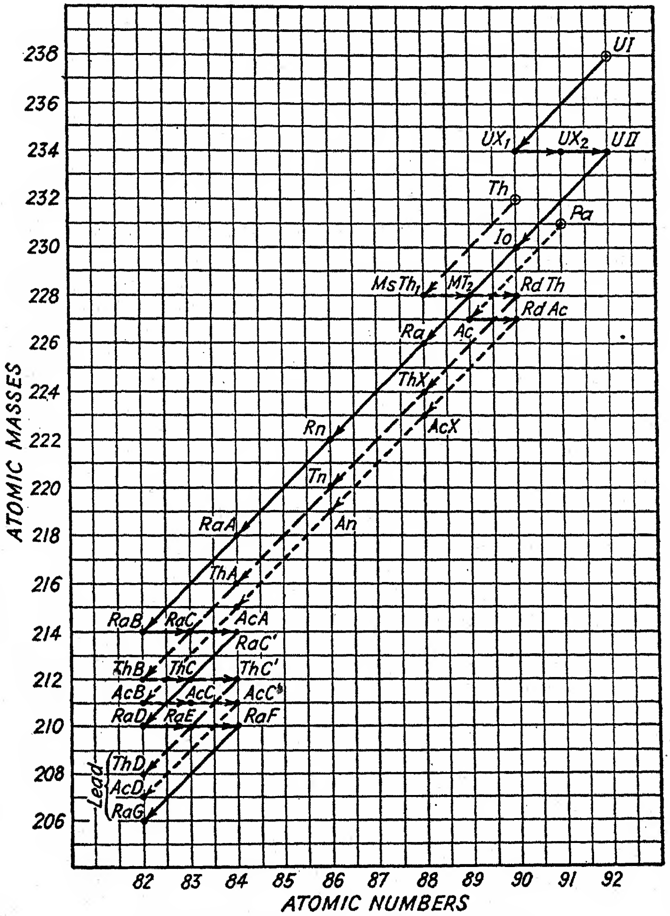


FIG. 625. Radioactive Transformation

series, the thorium series, and the actinium series. These are represented in Fig. 625, where an arrow downward and to the left indicates the ejection of an α-particle, and an arrow horizontal and to the right indicates the ejection of a β-particle.

Taking the uranium-radium series, for example, Uranium I (atomic mass 238, at. no. 92) emits an α-particle and becomes

UX_1 (at. mass 234, at. no. 90). UX_1 emits a β-particle and becomes UX_2 (at. mass 234, at. no. 91). This in turn emits a β-particle and becomes UII (at. mass 234, at. no. 92). UII emits an α-particle and becomes Ionium (at. mass 230, at. no. 90). Ionium emits an α-particle and becomes Radium (at. mass 226, at. no. 88), and so on. The uranium and thorium series are fully confirmed by experimental evidence, but the actinium series is not yet so well established.

It will be noted that in each case the end product is lead. The three kinds of lead have the same atomic number (82), but have atomic masses 206, 207, and 208, respectively. By means of the mass spectrograph, F. W. Aston has definitely distinguished these three kinds of lead.

636. Isotopes. Atoms which, like the leads above, have the same atomic number but different atomic weights have been given the name *isotopes* by Soddy. They prove to have the same chemical and physical properties except that of mass; hence they cannot be separated chemically.

In Fig. 625, points falling on the same vertical line represent isotopes. Thus, in addition to the three kinds of lead above mentioned, which are not radioactive elements, lead has four isotopes that are radioactive, viz.: Radium B(214), Thorium B(212), Actinium B(211), and Radium D(210).

637. Life of radioactive substances. The rate of disintegration of radioactive substances does not seem to be affected by ordinary chemical processes and physical conditions. Thus, temperatures from −186 to 2000°C, and pressures from that in the lowest vacuum up to 2000 atmospheres have no effect upon it.

The number of atoms disintegrating in a unit of time dn/dt at any time t is a constant fraction λ of the number of undisrupted atoms n at that time.

Expressed algebraically,

$$-\frac{dn}{dt} = \lambda n$$

where the (−) means that n decreases when t increases.

Integrating, and calling n_0 the number of undisrupted atoms when $t = 0$, we have the **law of disintegration:**

$$n = n_0 \epsilon^{-\lambda t}. \qquad (460)$$

Since for n to become zero, t must be ∞, it is customary in discussions of time of disintegration to use the *half-life period*, i.e., the time required for half the substance to disintegrate. To get the half-life period T, we make $n = n_0/2$ in Eq. (460) and solve for T.

$$T = \frac{2.303 \log_{10} 2}{\lambda} = \frac{0.6932}{\lambda}. \tag{461}$$

The half-life period is a measure of the stability of the radioactive element. It varies from 1.65×10^{10} years for thorium to 10^{-6} sec for radium C'. For radium, the most important radioactive substance, the half-life period is 1600 years.

Age of the earth. From the proportion of lead found in radioactive minerals, it is possible to estimate the age of the earth by means of the law of disintegration. By this method it appears that the age of the earth is not less than 1600 million years.*

638. Radon: the Curie. An atom of radium, on losing an α-particle, becomes an atom of gas which is itself radioactive. This gas is the "radium emanation" discovered by Rutherford. It has been found to belong to the group of chemically inert gases which includes neon, argon, etc., and is called *radon* (Rn), or niton (Nt). It has an atomic mass of 222 and an atomic number 86, and its half-life is 3.825 days.

Radon is very generally employed as the radioactive material in research and is what is commonly used in the treatment of cancer. It is pumped out of the vessel containing the radium and sealed in tiny, thin tubes of glass which are embedded in the flesh of the patient at the place to be treated.

The unit for measuring the quantity of radon is called the "curie." The **curie** is the amount of radon which is in equilibrium † with 1 gram of the element radium. One millicurie $\equiv$ 1/1000 curie.

639. Energy of radium. Pierre Curie and Laborde observed in 1903 that a radium salt maintained itself at a temperature several degrees (3° to 5°C) above the room temperature. Meyer and Hess have found the heat liberated by 1 gram of radium to be 132 gram-calories per hour, which is in good agreement with the value calculated from the energies of the α-, β-, and γ-rays separately.

* National Research Council, *Physics of the Earth*, Bul. No. 80, 1931, p. 454.

† I.e., the radon disintegrates at the same rate as that at which it is formed.

It is estimated that during the period required for a complete disintegration, 3 gm of radium (a five-cent piece weighs 5 gm) will liberate as much heat as a ton of coal. This great concentration of energy is of no use commercially, however, because radium is scarce and expensive.

640. Artificial radioactivity was first produced in January 1934 by F. Joliot and his wife, Irene Curie-Joliot.* The property was imparted to aluminum, boron, and magnesium by bombarding them with the α-particles from polonium.

By means of the cyclotron, Lawrence and his co-workers of the University of California at Berkeley have rendered practically all the elements radioactive by bombarding them with deuterons, the nuclei of heavy hydrogen.

641. Cosmic rays. From the time of Coulomb, the best insulated electroscopes were known to discharge gradually. The leakage seemed to be through the air surrounding the electroscope, and this required that ionization of this air must somehow be maintained. The question was, how? The most obvious source of ionization was the radioactive material which is widely distributed throughout the earth's crust. But this source was ruled out when V. F. Hess of the University of Innsbruck found (1913) that on balloon ascensions the ionization increased at higher altitudes, whereas if due to radioactive material in the earth it should have decreased. In 1936, Hess was recognized as the discoverer of cosmic rays and was awarded the Nobel prize.

Cosmic rays were so named after a suggestion by Mme. Curie that she suspected such a radiation disseminated throughout the universe. Subsequent investigation has tended to confirm this suspicion.

After 1925, Millikan and his co-workers determined the great penetrating power of the rays (6 ft of lead), and found the energy of some of them to be 6×10^9 electron-volts.

Kolhörster and Bothe greatly advanced the investigation by introducing the Geiger-Muller counter as a means of detecting the rays. This counter differs from that of Fig. 621 by having the

* It is interesting to note that artificial radioactivity was discovered by the daughter of the discoverers of radium. Mme. Curie-Joliot is the first child of a Nobel prize-winner to receive the prize also.

sharp-pointed electrode replaced by a fine wire stretched along the axis of the cylindrical electrode.

Cosmic ray "telescopes" were made by placing two of these counters parallel (Fig. 626). This determined the direction from which rays came, since (the counters being electrically in series) only rays coming in the plane of the two counters would ionize them both at once and thereby register. With sensitive devices of this kind, surveys have been widely made over land and sea by Arthur H. Compton and many others. It now seems definitely established that the intensity of the rays increases toward the magnetic poles, and that somewhat more rays come from the west than from the east.

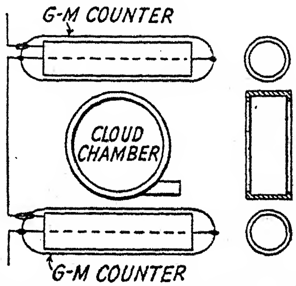


FIG. 626. Geiger-Muller Counters and Cloud Chamber

FIG. 627. Cosmic Ray Shower. (Courtesy Professor R. A. Millikan)

If cosmic rays were electromagnetic waves like light (photons), they would not be deflected by the earth's magnetic field; but if they consist of charged particles, they would be. The above facts

are in accord with the latter assumption; hence it is now generally believed that the rays consist in a large measure of charged particles. Compton adduces evidence tending to show that they are helium nuclei, + and − electrons, and protons. There is evidence that high energy photons also are present.

On passing through matter, such as a layer of lead, in a cloud chamber, cosmic rays often produce *showers* containing 20 or 30 electrons and positrons. The conclusion is drawn that the cosmic particles producing these showers must have masses approximately 150 times the mass of an electron. These heavy particles are called **mesotrons,** mesons, or barytrons. Figure 627 shows a shower.

The present evidence favors the hypothesis that cosmic rays originate in a uniform distribution throughout interstellar space. But their cause remains undetermined.

CHAPTER XLVIII

ATOMIC STRUCTURE—TRANSMUTATION OF ELEMENTS

642. Invisibility of atoms. From the theory of the microscope we know that two particles cannot be distinguished from each other if the distance between them is less than about one-third the wave length of the light that is used. Consequently, even with ultraviolet light ($\lambda = 3650$ Å), the smallest object that can be photographed (not seen) through a microscope is about 10^{-6} cm wide. Rutherford's experiment showed that the nucleus of an atom of gold cannot be more than 10^{-12} cm in diameter. Hence it appears hopeless for us ever to be able to see atoms directly by ordinary microscopic methods.

To get an idea of the structure of atoms we are therefore forced to adopt the methods of the detective and deduce from circumstantial evidence the make-up of the particular atom under consideration.

643. Early concepts. From very early times attempts were made to guess the shape of atoms from their behavior. Lucretius concluded, in the year 51 A.D., that substances that hurt the tongue must consist of sharp-pointed particles, whereas pleasant particles were smooth and round.

In 1815, Prout advanced the hypothesis that the atoms of all the elements were made up of hydrogen, since their atomic weights in terms of hydrogen were approximately whole numbers. This said nothing of the nature of hydrogen itself.

The kinetic theory of gases required only that the atoms or molecules should be tiny elastic spheres; but in 1859 the discovery by Kirchhoff and Bunsen that each element emitted a characteristic spectrum showed that something more was necessary. Since light was known to have wave properties, this discovery seemed to make it necessary that an atom should oscillate or contain some type of oscillator in order that it might emit light of proper

wave lengths. However, no type of vibrating atom model achieved any success in accounting for spectra.

644. The electron. The discovery of the electron in 1897 by J. J. Thomson and his co-workers is perhaps the greatest achievement in the whole history of physical science. Of it Karl T. Compton has said: ". . . no . . . instance has been so dramatic as the discovery of the electron, the tiniest thing in the universe, which within one generation has transformed a stagnant science of physics, a descriptive science of chemistry and a sterile science of astronomy into dynamically developing sciences fraught with intellectual adventure, interrelating interpretations and practical values." *

The electron has been found to be always the same, regardless of its source. The latest value of the quantity, or charge, of electricity represented by one electron is:

$$1 \text{ electron} = (4.8025 \pm 0.0010) \times 10^{-10} \text{ statcoulomb.}$$

This was determined by R. T. Birge of University of California at Berkeley, as the most probable value from present experimental data.†

The term **electron** almost always refers to the − charge of the above amount. The equal positive charge is called a **positron.**

Since the ratio of the charge to the mass was already accurately known (Sec. 505), we have:

$$\text{rest mass of 1 electron} = (9.1066 \pm 0.0032) \times 10^{-28} \text{ gram.}$$

The electron has been found to possess the properties of a wave as well as those of a particle.‡

645. The proton. By his method of positive ray analysis,§ J. J. Thomson and F. W. Aston found that when a hydrogen atom has been ionized by the loss of its one electron, the remaining + ion has 1700 times the mass of the electron. (Birge's more recent value is 1836.)

* *Science* (Jan. 8, 1937).
† R. T. Birge, *Reviews of Modern Physics*, **4**, 233, 1941.
‡ C. J. Davisson, "The Wave Properties of Electrons," *Science* (June 27, 1930).
§ Glazebrook, *Dictionary of Applied Physics* (New York, The Macmillan Company), II, p. 603.

Since the atoms of hydrogen were neutral to begin with, and a − electron was removed, the + charge of the remaining ion must have the magnitude of the charge of 1 electron. This positively charged nucleus of the hydrogen atom is called a *proton*. Its mass ($1.672 = 10^{-24}$ gm) is practically that of the hydrogen atom, 1.673×10^{-24} gram.

646. The alpha-particle. From their mass and charge, it seemed probable that α-particles were ionized atoms of helium. By a brilliant experiment, Rutherford and Royds showed beyond peradventure that this was the case: An alpha-particle is what remains after a helium atom has lost its two electrons.

$$\text{Mass of } \alpha\text{-particle} = 4 \text{ times mass of proton}$$
$$\text{Charge of } \alpha\text{-particle} = +2e.$$

Figure 631 shows the trails of water droplets that have condensed on ions along the "track," or path, of an α-particle in a Wilson cloud chamber. Because of the great energy and charge of α-particles, their tracks are wider (more droplets) than those of electrons or protons.

The abrupt change of direction of the track at the extreme left in the figure shows where an α-particle has collided with an atom of nitrogen, ejecting a proton and leaving an atom of oxygen. The wide trail below the fork is that of the α-particle; the narrow trail forming the right branch is the path of the proton; and the wide left branch is that of the atom of oxygen, which farther on encounters another atom. (See Sec. 654.)

647. The nucleus. Rutherford, Geiger, and Marsden bombarded thin films of gold, silver, copper, and platinum with α-particles from radium (*B* and *C*).* Most of these α-particles went through undeflected, leaving no detectable holes; but about 1 in 500 or 1 in 1000 was deflected from its original direction, or "scattered," through angles varying from 5° to 150°.

From these data three conclusions were drawn:

1. That even dense matter like gold has a very open-work structure, since only 1/5 of 1% of the α-particles were deflected.

* Rutherford, Chadwick and Ellis, *op. cit.*, p. 195.

2. That the gold and other atoms must be built up around a small nucleus having a + charge and most of the mass of the atom; and that the positively charged α-particles were repelled according to Coulomb's law when they passed near the nucleus.

3. From the angles of deviation, Chadwick calculated that the nucleus must have a net + charge equal to its atomic number, thus confirming the conclusion drawn by Moseley from his work on x-ray spectra (Sec. 615).

From similar experiments on the scattering of α-particles, they conclude that the diameter of the gold nucleus cannot be over 6.4×10^{-12} cm, and that of copper 2.4×10^{-12} cm; while similar scattering by hydrogen indicates the diameter of its nucleus, the proton, to be not more than 3.0×10^{-13} cm.

648. The Rutherford-Bohr atom. In 1911, Lord Rutherford proposed an atomic model consisting of a small nucleus having practically all the mass of the atom and a + charge equal to its atomic number times *e*. Around this nucleus he supposed − electrons to revolve like planets around the sun. In a neutral atom the number of these electrons must obviously equal the atomic number also.

According to the principles of classical physics, atoms of this type should emit spectra under ordinary conditions; and in doing so, the electrons should spiral into the nucleus, which means that the atom would collapse. But it is well known that neither of these things takes place. Hence, either the model was wrong or the laws of classical physics did not hold within the atom.

Adopting the nuclear atom of Rutherford, Niels Bohr, professor of physical chemistry at the University of Copenhagen, cast aside the sterile stipulations of classic theory and espoused the questionable quantum theory of Max Planck (Sec. 310). Bohr's assumptions were as follows:

1. In a neutral atom under ordinary conditions, the electrons revolve in orbits in a "steady state," * i.e., with constant speed, without radiating energy. This is contrary to Maxwell's electromagnetic theory.

2. Energy is radiated only when an electron "jumps" from an orbit farther out to one nearer to the nucleus; and the energy

* Called also "stationary state," or "energy level."

radiated is the difference between the energies that it possessed in these two steady states, as given by Eq. (449), which is known as **Bohr's frequency relation:**

$$h\nu = W_2 - W_1.$$

3. It is possible for an electron in the steady state to occupy only certain definite orbits, and these are "selected" by (i.e., must satisfy) the relation:

$$2\pi mvr = nh \tag{462}$$

where m is the mass of the electron;
v is the linear speed of the electron in its orbit;
r is the radius of the orbit;
n is a whole number: 1, 2, 3, $\cdots$ n; and
h is Planck's universal constant.

Bohr's reason for making these assumptions was that they give the answer, already known, viz., Balmer's formula. His combination of classic physics, quantum theory, and arbitrary assumptions leads to the relation: *

$$W_n = -\frac{2\pi^2 me^4Z^2}{n^2h^2}, \quad n = 1, 2, 3, \cdots \tag{463}$$

as the total energy of an electron in the n^{th} energy level, or orbit, where Z is the atomic number of the element.

Hence, in falling from an outer orbit n_2 to an inner orbit n_1 the energy given out is:

$$h\nu = W_{n_2} - W_{n_1} = \frac{2\pi^2 me^4Z^2}{h^2}\left(\frac{1}{n_1{}^2} - \frac{1}{n_2{}^2}\right). \tag{464}$$

By comparison with Eq. (448), it is seen that the *Rydberg constant* is:

$$R = \frac{2\pi^2 me^4Z^2}{ch^3}.$$

Computations on the basis of Bohr's theory give the Rydberg constant and the spectral lines of **"hydrogen-like"** atoms, i.e., atoms having but one electron outside the nucleus, with very satisfactory accuracy. The theory also predicted the spectral

* Ruark and Urey, *Atoms, Molecules, and Quanta* (New York, McGraw-Hill Book Company, 1930), p. 68.

series discovered by Lyman, Paschen, Brackett, and Pfund already mentioned in Sec. 616. But even with elliptic orbits and the relativity corrections of Sommerfeld, the theory did not predict accurately the energy of the preferred orbits. And it failed entirely for atoms having more than one electron outside the nucleus, for this involved the "problem of three bodies," which is still not completely solved.

Bohr's atom model, like its predecessors, has ceased to be regarded as giving a correct geometrical picture of atomic structure. But it was a distinct forward step for it showed that the Rydberg constant is related to the other fundamental constants; it served to establish the concept of "energy states," or levels (Fig. 601); and it was the first model to achieve any success in the explanation of spectra.

649. Atoms heavier than hydrogen. From chemistry it will be recalled that He(2), Ne(10), Ar(18), Kr(36), and Xe(54) * are inert elements, forming no compounds under ordinary circumstances. It has been shown also that an atom of each element has outside its nucleus a number of electrons equal to its atomic number.

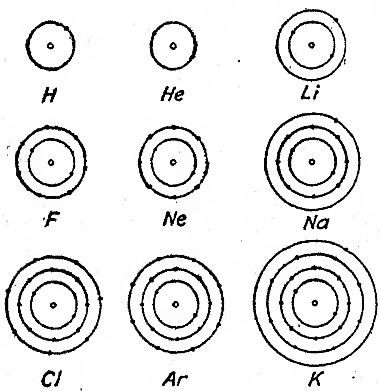


FIG. 628. Symbolic Diagrams of Electron Shells

From such facts, Sir J. J. Thomson suggested that these electrons may be arranged in layers, or "shells," and that **a complete (or completely filled) shell is a very stable arrangement,** so that atoms having complete shells are therefore inert.

According to this idea, helium has 1 complete shell containing 2 electrons. Neon has 2 complete shells, the inner one containing 2 electrons and the outer containing 8 electrons. Argon has 3 complete shells, the inner containing 2 electrons, the next containing 8 electrons, and the third containing 8 electrons; and so on (Fig. 628).

* Numbers in parentheses are the atomic numbers of these elements.

Atoms having atomic numbers between helium and neon will have a first shell of 2 electrons with 1 or more additional electrons in the second shell. Atoms between neon and argon will have a first shell of 2 electrons, a second shell of 8 electrons, and 1 or more electrons in the third shell, and so on.*

Atoms which do not have the outer shell complete might be expected to be active, since they could readily lose or gain one or more electrons. Consider the following table:

Element	Atomic Number	Number of Shell and Number of Electrons in Shell			
		I	II	III	IV
H	1	1			
He	2	2	(Complete)		
Li	3	2	1		
F	9	2	7		
Ne	10	2	8	(Complete)	
Na	11	2	8	1	
Cl	17	2	8	7	
Ar	18	2	8	8	(Complete)
K	19	2	8	8	1

Here it will be seen that hydrogen lacks 1 electron of having the first shell complete. Therefore it might either lose this electron or combine with some element requiring 1 more electron to fill out a complete shell. It does both. In electrolysis, it loses its electron and with a water molecule forms a "hydronium ion" $(H_3O)^+$. It makes with another hydrogen atom the fairly stable hydrogen molecule H_2; and combines vigorously with chlorine, which lacks 1 electron in its third shell, to form HCl.

Li, Na, and K each have 1 electron in excess of a complete shell; hence they can readily lose this outer, or valence, electron. They are active and have similar properties. Their ions in solution are electropositive, indicating that they do lose this extra electron.

F and Cl each lack 1 electron of having a complete shell. They also are very active chemically and have similar properties. Their ions in solution are electronegative, indicating that they readily take up another electron.

* For complete table, see Richtmyer, *Introduction to Modern Physics* (New York, McGraw-Hill Book Company, 1928), p. 572.

Electrons are more easily removed from the outer shell than from the inner shells. Thus, in the case of K, the work required to remove the outermost, or valence, electron is 4.32 electron-volts, but 19 electron-volts are necessary to remove the first electron from shell III, and 48 electron-volts to remove one from shell II.

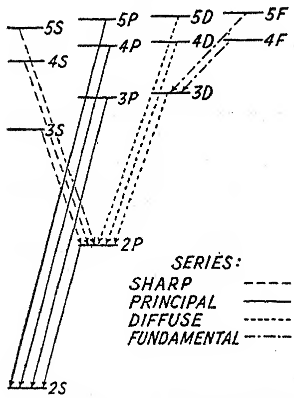


FIG. 629. Energy Levels for Lithium

The idea of energy levels, which was so successful in accounting for the spectra of hydrogen, has been extended to the more complex elements. In these cases, if an electron might jump from any level of higher energy to any level of lower energy, more spectral lines would be predicted than show up in spectrograms. But comparatively simple rules for selecting the proper energy levels have been formulated which enable us to predict just those lines that are actually found.

Figure 629 shows a simplified energy-level diagram for lithium. The selection rules require an electron to jump from levels of higher to those of lower energy in adjacent columns.

650. The neutron. In 1930, Bothe and Becker of Giessen observed that when beryllium is bombarded with α-particles from polonium, a very penetrating radiation is emitted which they assumed to be γ-rays.

But in June 1932, James Chadwick at the Cavendish laboratory showed from energy considerations that the radiation they had produced must be a particle without charge, having the mass 1.0067 of atomic mass units. To this particle he gave the name "neutron." A *neutron* is a neutral particle having the mass of a hydrogen nucleus. It has been given the symbol $({}_0n^1)$.

The reaction of Bothe and Becker's experiment is written:

$$ {}_2\mathrm{He}^4 + {}_4\mathrm{Be}^9 \rightarrow {}_6\mathrm{C}^{12} + {}_0n^1 + E \qquad (465) $$

where **the subscript represents the nuclear charge of the atom and the superscript, the atomic mass.** E is the energy necessary to balance the equation.

Having no charge and no attendant electrons, neutrons do not influence other particles except when they make a direct hit; i.e.,

they do not ionize by electric forces as do protons, deuterons, and α- and β-rays. They themselves consequently do not make tracks in cloud chambers, but they knock protons out of hydrogen compounds like paraffin and α-particles out of nitrogen, and these latter particles do produce tracks in the cloud chamber. Also, since they have no charge, they are not so readily deflected as are charged particles; hence they are more penetrating than the hardest γ-rays.

Whether neutrons are a very close combination of a proton and an electron, or an independent entity, has not yet been definitely determined; but they can be knocked out of the nuclei of other elements such as Li and Al, and must therefore be considered one of the building blocks of nature.

651. The positron. Dirac and other mathematical physicists had suspected from theoretical considerations that a positive particle having the same mass as the electron existed, but it was not until August 2, 1932, that one was recognized experimentally. The discovery was made by Dr. Carl D. Anderson of the California Institute of Technology.

A vertical cloud chamber had been placed between the poles of a strong electromagnet for the study of cosmic rays. One Geiger-Mueller counter was placed just above the cloud chamber and another just below it, so that a cosmic ray that traversed both counters would actuate the camera and automatically take the photograph of the corresponding ray tracks in the cloud chamber.

The purpose of the magnetic field was to enable the energy of the trackmaking particles to be determined. Since a moving charge of electricity is equivalent to a current (Rowland's Experiment, Sec. 408B), the path of the charge would be curved by the magnetic field in accordance with Fleming's motor rule. From the intensity of the magnetic field and the curvature of the path, with other known data, the energy of the particle can be computed. The greater the energy of a particle the less the curvature.

Figure 630 is a reproduction of the photograph by means of which the positron was discovered. Across the diameter of the cloud chamber is a lead plate 6 mm thick. The direction of the magnetic field was perpendicular and into the paper.

The beaded appearance of the track indicates that it was made

by a particle of the mass and charge of an electron. The curvature below the plate corresponds to an energy of 63 million electron-volts (mev), while the curvature above the plate corresponds to 23 mev. If the particle were a + electron, the direction of curvature shows that it must have come downward. But it is inconceivable that an electron would gain energy in passing through 6 mm of lead. On the other hand, if the track were due to a + electron of the same mass as the − electron and moving upward, the curvature is in the right direction, the energy change is right, and the appearance of the path is proper. Anderson therefore concluded that the particle that produced the ionization was such a positive particle, and gave it the name "positron."

FIG. 630. First Photograph of Positron Track. (Courtesy of Dr. Carl D. Anderson)

A **positron** is a particle of positive electricity having the same mass and charge as the familiar negative electron.

Subsequently it has been found that positrons are emitted by materials that have been rendered artificially radioactive, and by the impact of λ-rays from thorium C' on lead. It appears that a photon of a billion electron-volts can produce showers of electrons both + and −; but showers are also occasionally produced by photons of 3 or 4 mev.

Figure 627 is a photograph of such a shower, presumably from a cosmic-ray photon striking lead at A. The two tracks downward to the right are made by positrons of 145 and 38 mev, respectively, while the three downward and to the left are due to negative electrons of 104, 65, and 28 mev, respectively.*

Positrons apparently play no part in conduction. They combine promptly with electrons, probably producing photons of radiant energy.

652. Hydrogen isotopes. At one time, the atomic mass of hydrogen as determined chemically (1.008) differed from the value

* Millikan, *Electrons* (+ and −), etc. (University of Chicago Press, 1935), p. 345.

obtained by means of the mass spectrograph (1.0078). From this circumstance, mathematical physicists suspected that ordinary hydrogen was a mixture of a light isotope ($_1H^1$) and a heavier one.

If that were true, it was to be expected that these two kinds of hydrogen in the liquid state would have different boiling points and hence could be separated or concentrated by fractional distillation.,

Accordingly, **in the memorable year, 1932,** Urey, Murphy, and Brickwedde of Columbia University and the Bureau of Standards succeeded by that method in getting a sufficient concentration of the heavy isotope to enable it to be definitely, though faintly, recognized in spectrograms of the residual mixture.

This heavy isotope is called **deuterium** and has been given the symbol $_1H^2$ or $_1D^2$, since its atomic mass is by computation 2 and by the mass spectrograph, 2.0136.

A *deuteron* is the nucleus of an atom of deuterium. Its symbol is also generally written, $_1H^2$. A deuteron is believed to consist of 1 proton and 1 neutron.

Heavy water ($H_2{}^2O$), or deuterium oxide, is now prepared commercially by electrolyzing ordinary water. Recently, a third isotope of hydrogen ($_1H^3$) has been produced in very small amounts by electric discharge in deuterium. Since there are also three isotopes of oxygen ($_8O^{16}$, $_8O^{17}$, $_8O^{18}$), it is possible to have 18 different kinds of pure water.

653. Composition of the nucleus. Before the discovery of the neutron and the positron, the nucleus was thought to consist of a number of protons equal to the atomic mass, these protons being held together by such a number of electrons that the net + charge of the nucleus was equal to its atomic number. Thus the helium nucleus, or α-particle, was believed to consist of 4 protons held together by 2 electrons; and the nucleus of nitrogen was supposed to consist of 14 protons held together by 7 electrons, giving a net + charge equal to its atomic number 7.

But with the discovery of the neutron all this changed.

It now appears that the nuclei of atoms consist of protons and neutrons only—no electrons. The number of protons is always equal to the atomic number of the element; the rest of the atomic mass is made up by the proper number of neutrons. Isotopes have

the same number of protons, but different numbers of neutrons. Some examples are given in the following table.

Nucleus	Symbol	No. of Protons (Atomic Number, Z)	No. of Neutrons
Hydrogen (1)	$_1H^1$ (proton)	1	0
Hydrogen (2)	$_1H^2$ (deuteron)	1	1
Hydrogen (3)	$_1H^3$	1	2
Helium	$_2H^4$ (alpha)	2	2
Lithium (1)	$_3Li^6$	3	3
Lithium (2)	$_3Li^7$	3	4
Sodium (1)	$_{11}Na^{23}$	11	12
Sodium (2)	$_{11}Na^{24}$	11	13
Uranium	$_{92}U^{238}$	92	146

The α-particle $_2H^4$ is apparently a very stable combination of 2 protons and 2 neutrons, for it is ejected intact from the nuclei of natural radioactive substances. The electron which constitutes a beta-particle is believed to come from the breakdown of a neutron into a proton and an electron at the instant of the explosion, only the electron being expelled.

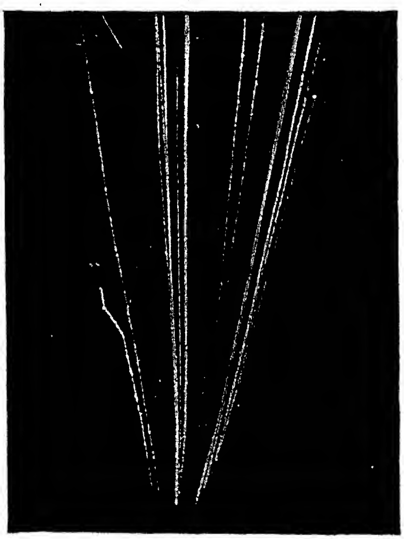


Fig. 631. First Artificial Transmutation of Elements. (Courtesy of Dr. P. S. M. Blackett)

654. Nuclear disintegration. For nearly a century after the atomic theory was advanced by the English schoolteacher, John Dalton, in 1802, chemists and physicists believed that the atoms of the elements were the ultimate, indivisible particles of matter. This idea was upset when natural radioactivity showed the heavier atoms spontaneously breaking up into α-particles, electrons, photons, and a residual lighter atom.

The first instance of artificial disintegration was recognized by Rutherford in 1919 when he knocked protons out of nitrogen nuclei by bombarding the latter with α-particles from radium C.

He estimated that about 20 such cases occurred when 1,000,000 alphas were employed. Dr. P. S. M. Blackett of the Cavendish laboratory made 23,000 photographs of the tracks of alphas in a cloud chamber filled with nitrogen. Among these, 8 showed encounters similar to that shown in Fig. 631, which is one of Blackett's photographs.

At the left of this picture one of the tracks is seen to be forked. The stem of the fork is the track of the alpha; the thin right branch is the track of the proton; and the thick left branch is the track of an oxygen nucleus resulting from the encounter.

Using the symbolism of the two preceding sections, this nuclear episode may be expressed as follows:

$$ {}_{7}\mathrm{N}^{14} + {}_{2}\mathrm{He}^{4} \rightarrow {}_{8}\mathrm{O}^{17} + {}_{1}\mathrm{H}^{1} + E. \tag{466}$$

This is at once a disintegration and a transmutation, the first of either to be brought about artificially. It may be described in words by saying that an α-particle ${}_{2}\mathrm{H}^{4}$ combined with a nitrogen nucleus ${}_{7}\mathrm{N}^{14}$ and produced a nucleus of an isotope of oxygen ${}_{8}\mathrm{O}^{17}$ and a proton, or hydrogen nucleus, ${}_{1}\mathrm{H}^{1}$.

In this experiment, the physicist had realized the dream of the alchemist, except in its alluring commercial aspect.

655. The cyclotron. The cyclotron, devised by E. O. Lawrence and M. S. Livingston of the University of California at Berkeley,* is an exceedingly effective "atom-smasher," based on a new principle which does not require such high voltages as are developed by Van de Graaff's machine or the a-c rectifying system of Cockroft and Walton.

It consists of an enormous magnet (Fig. 632), between the poles of which is placed the device shown in Fig. 633. The latter consists of two shallow semicircular boxes X and Y, an open top view of which is exhibited. The diametral wall of each of these boxes, called the "dees," has a narrow central slot along nearly its entire length. The two boxes are insulated from each other and form the two plates of a condenser, the dielectric being the rarefied gas in the channel MN between the two diametral walls. These semicircular boxes may be charged alternately + and − by means of a vacuum tube oscillator (not shown) connected at OO. The entire circular condenser is enclosed in another flat, circular,

* E. O. Lawrence and M. S. Livingston, in the *Physical Review* (April 1, 1932).

air-tight box, or vacuum chamber, and this assembly is placed between the poles of the magnet.

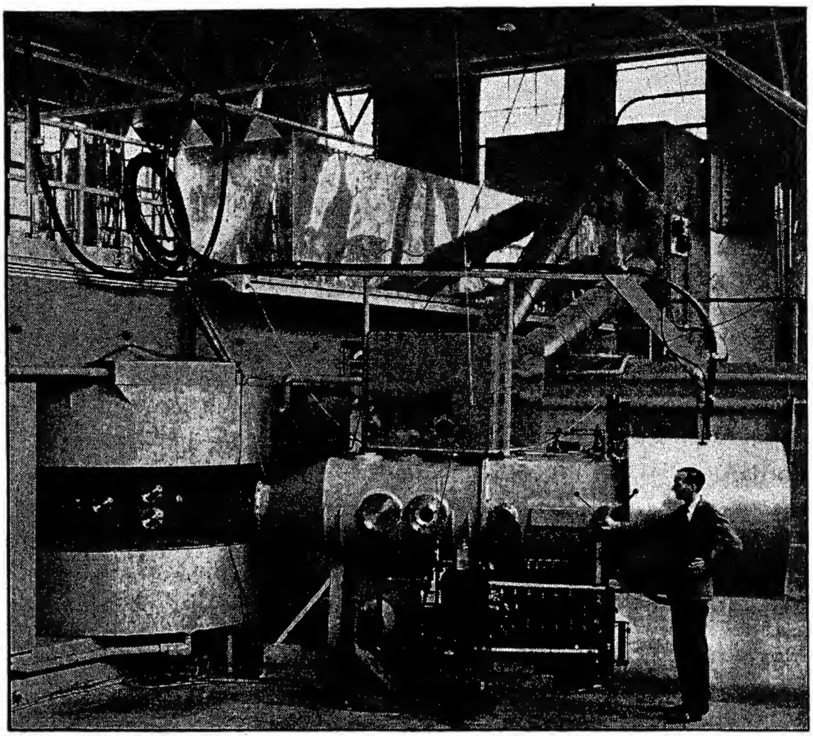


FIG. 632. Lawrence's 60 Inch Cyclotron. (Courtesy Dr. E. O. Lawrence)

If the projectile is to be a deuteron, the vacuum chamber is exhausted to a pressure of about 10^{-6} cm of Hg, and a small quantity of deuterium is then admitted. Electrons emitted from two small, hot spiral filaments at F are accelerated between these filaments sufficiently to ionize the deuterium.

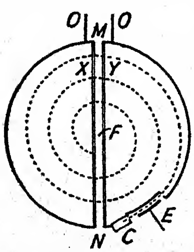


FIG. 633. Sketch of the Dees

Suppose that at the instant of ionization of a deuteron, the side X is $-$ and Y is $+$. The deuteron is driven into half-cylinder X by the electrostatic field. Since a moving charge has the same properties as a current, the path of the ion will be bent into a semicircle by the magnetic field, whose intensity H is about 18,000 oersteds, in accord with Fleming's motor rule (Sec. 394).

The centripetal force is due to the magnetic field and has the value Hev.

Hence,

$$Hev = \frac{mv^2}{r}$$

and the time required to traverse one semicircle is:

$$t = \frac{\pi r}{v}$$

whence

$$t = \frac{\pi m}{He} \tag{467}$$

which is independent of the radius. Hence all the semicircles are traversed in equal times. Consequently, it is possible for the period of an oscillator to be adjusted so that the polarity of the half-cylinders is reversed each time the ion crosses the diametral channel.

If the potential difference between opposite sides of the channel (condenser) is 50,000 volts, the energy of the deuteron is increased by 50,000 electron-volts each time it passes across the channel. If it makes 100 half-circles, its final energy is 5,000,000 ev (electron-volts).

The material to be bombarded is placed in a container at C; or the deuterons may be fired out into the open air through a window in the side of the case, covered with metal just thick enough to withstand atmospheric pressure and maintain the vacuum. E is an electrode to enable the beam of ions to be directed upon the slit by an electrostatic field.

The huge magnet is about 9 ft high and has a mass of 65 tons. Nine tons of heavy copper strip were required to wind the magnet coils. The diameter of the magnet poles is 45 in.

In a cyclotron operating at a few mev, we have a means of producing more neutrons than could be produced by bombarding beryllium with the α-particles of all the radium of any country in the world if collected all at one place.

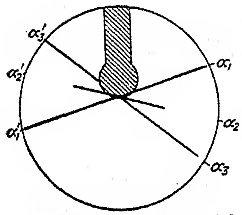
656. Transmutation of the elements. Up to 1932, nuclear disintegration and its accompanying transmutation had been brought about only by the impact of α-particles from natural radioactive substances.

In that year, Cockroft and Walton, at the Cavendish Laboratory, Cambridge, "the mother laboratory of modern physics," employing a number of kenotrons in series, succeeded in producing protons with energies of 150,000 electron-volts. With these they bombarded lithium and produced the **first transmutation by means of artificially produced projectiles (protons)**, as follows:

$$_1H^1 + {}_3Li^7 \rightarrow {}_2H^4 + {}_2H^4 + E. \qquad (468)$$

The hydrogen nucleus (proton) combined with the lithium nucleus to produce two helium nuclei (α-particles).

These two helium nuclei are shot out in opposite directions with the great kinetic energy of 8.53×10^6 electron-volts each. Figure 634 shows three such pairs. This energy, less the initial kinetic energy of 150,000 ev of the proton, is represented in Eq. (468) by E. From the magnitude of the energies of the alphas, it is seen that the reaction is in the nature of an explosion.



After Walton & Dee (Cavendish Lab'y)

FIG. 634. Transmutation of Lithium into Helium

The question of whence came this enormous energy exhibited by the alphas is discussed in Sec. 657.

So many transmutations have been produced since 1932 that it is rather difficult to make a choice for illustration.* The following one, however, is of special interest.

Using the cyclotron, Lawrence bombarded a piece of rock salt with a stream of deuterons of 2.15 mev in the air for an hour. After a preliminary period of about 20 hours in which its radiations were complex and uncertain, its decay became exponential like that of natural radioactive substances and continued, the half-life period being 15.5 hours.

The reaction was as follows:

$$_{11}Na^{23} + {}_1H^2 \rightarrow {}_{11}Na^{24} + {}_1H^1 \qquad (469)$$

where $_{11}N^{24}$ is radioactive and may be obtained in large quantities. It is called **radio sodium.**

* J. R. Dunning and George B. Pegram of Columbia University have transmuted platinum into gold by bombarding the former with α-particles from radium. (*Science News Letter*, Nov. 17, 1939.)

Radio sodium seems to break down according to the following relation:

$$_{11}Na^{24} \rightarrow {}_{12}Mg^{24} + {}_{-1}e^0 + \gamma + E. \tag{470}$$

That is, it yields ordinary Mg, an electron $_{-1}e^0$, and gamma radiation. E is the amount of energy necessary to balance the equation, i.e., the kinetic energy of the electron and the Mg atom.

Since both sodium and magnesium are common ingredients of food, radio sodium would seem to be admirably suited for internal treatments with γ-rays, in view of its short half-life period.

Surprising also is the fact that radio sodium may be prepared by three other transmutations:

$$_{11}Na^{23} + {}_0n^1 \rightarrow {}_{11}Na^{24}$$
$$_{13}Al^{27} + {}_0n^1 \rightarrow {}_{11}Na^{24} + {}_2He^4 \qquad \text{(due to Fermi)}$$
$$_{12}Mg^{24} + {}_0n^1 \rightarrow {}_{11}Na^{24} + {}_1H^1.$$

657. Transmutation of mass and energy. One of the conclusions from Einstein's theory of relativity (Sec. 628) was that mass should be changeable into energy or energy into mass according to the relation:

$$E = mc^2$$

where m is the mass in question;
c is the velocity of light;
E is the energy equivalent to mass m.

This prediction of the relativity theory is verified in a highly satisfactory way by the reaction of Eq. (468). Taking atomic masses in terms of $\frac{1}{16}$ of the mass of an oxygen atom as a unit, to which the name *dalton* has been given, we have the following values:

Mass of Atom	Deduct Electrons	Mass of Nucleus
e = 0.00055		
H = 1.0081	1	1.0075
He = 4.00216	2	4.0011
Li = 7.0148	3	7.0132

Writing these masses for the nuclei in Eq. (468), we have:

$$_1H^1 + {}_3Li^7 = {}_2He^4 + {}_2He^4 + E$$
$$1.0075 + 7.0132 = 2(4.0011) + E$$
$$0.0185 = E.$$

Hence, if the laws of conservation of mass and of energy hold true, the energy E must equal 0.0185 dalton.

It may be shown that

$$1 \text{ dalton} = 927 \times 10^6 \text{ electron-volts.}$$

Therefore

$$0.0185 \text{ dalton} = 0.0185 \times 927 \times 10^2 = 17.15 \text{ mev.}$$

The observed value of the energy of the two alphas in Sec. 656 was 17.06 mev.

The calculated and experimental values are in excellent agreement and offer strong confirmation of the theory.

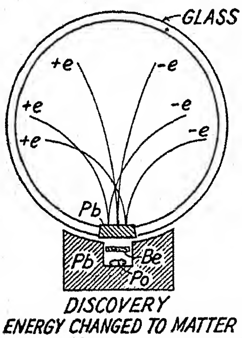


FIG. 635. Transformation of Energy into Matter

In the foregoing we have an example of the transformation of matter into energy. The reverse effect—the transformation of energy into matter—was discovered by I. Curie and F. Joliot in an experiment described in Fig. 635.

Alpha-rays from polonium Po impinge upon a sheet of beryllium Be, causing the emission of γ-rays of 5×10^6 electron-volts energy. These γ-rays pass through the lead plug in the side of the glass cloud chamber, and electrons and positrons in pairs spring from the surface of the lead within the cloud chamber, as shown.

Figure 636 is a cloud chamber photograph by Dr. W. A. Fowler of California Institute of Technology. In this instance a photon of γ-radiation (which makes no track) has accidentally been caught at the instant when it was transformed into an electron and a positron, as it passed through the gas of the cloud chamber.

The rest mass of an electron (either K or $+$) is 0.00055 dalton, and by Einstein's relation ($E = mc^2$) the energy equivalent is:

$$0.00055 \times 927 \times 10^6 = 510{,}000 \text{ ev.}$$

Hence, to produce two of these electrons would require a photon of at least 1,020,000 ev.

The γ-rays from thorium C'' have energy of 2.65×10^6 ev.* So if these are used, the two electrons would be produced according to the following relation:

$$L \rightarrow {}_1e^0 + {}_{-1}e^0 + E \tag{471}$$

where L is the photon.

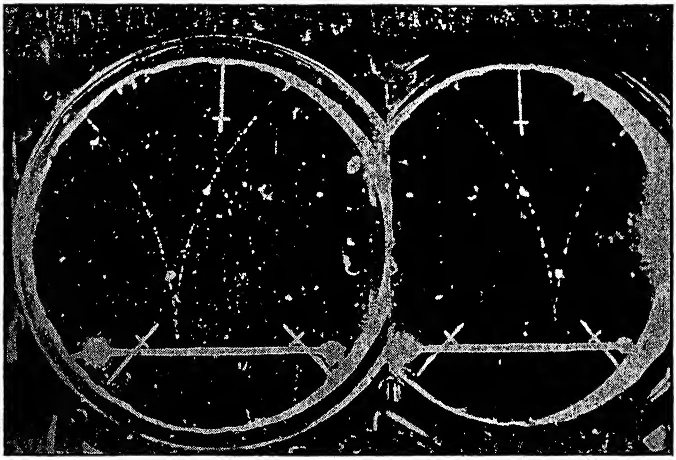


FIG. 636. Electron Pair. (Courtesy of Dr. W. A. Fowler)

The residual energy $E = (2.65 - 1.02) \times 10^6 = 1.63$ mev would be divided equally between the electron and the positron, giving to each a kinetic energy of 815,000 ev, as is shown by the curvature of their trails.

Measurements of actual photographs show these computations are in accord with the facts.

Positrons and electrons in pairs are readily obtained by irradiating heavy elements such as lead with γ-rays of great energy.

658. Waves and particles. In the study of light, we saw that in Young's experiment light had the property of waves, whereas the photoelectric effect can be explained only by Einstein's assumption that light consists of discrete packages $h\nu$, or particles. Einstein's theory of relativity predicted that light quanta should be deflected by a strong gravitational field like that of the sun; i.e., that light had gravitational mass—and this was verified by experiment.

* These γ-rays have the greatest energy of such rays from any natural radioactive substance.

Since light thus appeared to have mass and discontinuous structure like matter, it occurred to Louis de Broglie in 1924 that matter also might have a dual nature and possess the property of waves. This theory of material waves associates a definite wave length with every particle of matter according to the relation:

$$\lambda = \frac{h}{mv} \qquad (472)$$

where λ is the wave length associated with the particle;
h is Planck's constant;
m is the mass of the particle; and
v is the velocity of the particle.

In 1927, this theory was confirmed by an experiment of Davisson and Germer in the Bell Telephone Laboratories. They found that electrons are diffracted by a crystal of nickel, just as are x-rays. Their observations of angles and intensities were made on a type of spectrometer and do not lend themselves to brief description.

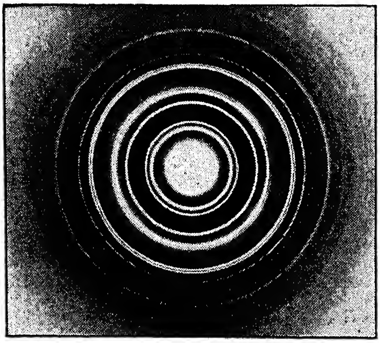


FIG. 637. Diffraction of Electrons by Aluminum. (Courtesy of Dr. L. H. Germer)

Shortly afterward, however, G. P. Thomson (the son of Sir J. J. Thomson) passed a narrow circular beam of electrons through a sheet of gold foil about 10^{-6} mm thick, let the emerging diffracted electrons fall upon a photographic plate parallel to the gold foil, and secured a picture similar to Fig. 637. It is seen to be of the same type as Fig. 638, obtained by Webster Richardson by passing a narrow beam of x-rays through a similar thin sheet in which the crystals were arranged in completely random fashion.

The wave lengths computed from the experiment of Davisson and Germer as well as from Thomson's were in good agreement with the values predicted by Eq. (472).

This duality in the nature of matter seems now a natural inference if energy (radiation) and matter are mutually convertible;

but this latter fact was not established until nearly ten years after de Broglie's brilliant concept.

The theory of material waves at once puts Bohr's arbitrary criterion for preferred orbits on a rational basis. If we are to have periodic motion in a closed orbit, the length of the orbit must be an integral multiple n of the wave length λ of the material wave. If r is the radius of the orbit, then

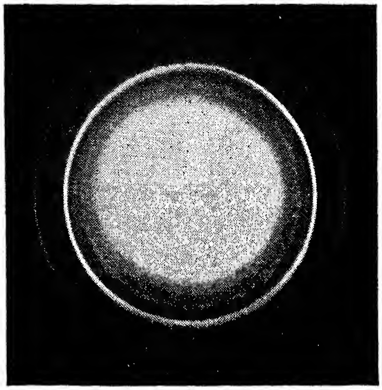


FIG. 638. Diffraction of X-Rays by Aluminum. (Courtesy Professor W. Richardson)

$$2\pi r = n\lambda = n\frac{mv}{h}\cdot \quad \text{by Eq. (472)}$$

Therefore,

$$mvr = n\frac{2\pi}{h}$$

which is Bohr's criterion (Sec. 648).

659. Wave mechanics. In 1925, Erwin Schroedinger of the University of Berlin initiated the principle of building up the science of mechanics from the idea of material waves rather than from that of material particles. This new *wave mechanics* has the same relation to the classical mechanics of Newton that physical optics has to geometrical optics.

Just as the electromagnetic theory of light has been said to be "Maxwell's equations," so Schroedinger's theory for a single particle having one-dimensional motion may best be described by his wave equation:

$$\frac{d^2\Psi}{dx^2} + \frac{8\pi^2 m}{h^2}(E - V)\Psi = 0. \qquad (473)$$

This equation represents the motion of a particle of mass m along the x-axis. E and V are the total energy and the potential energy, respectively, of the particle, and Ψ is a complex function whose absolute square, according to Schroedinger, represents the probability that m shall be at a given point at a certain time.

Just a few months after the publication of Schroedinger's

theory, Werner Heisenberg of Leipsic University brought out a new science of motion which he called **quantum mechanics.**

Whereas Schroedinger and de Broglie adhered to the methods of classical physics, Heisenberg began with the assumption that all inherited concepts of classical physics should be abandoned and the new science built up entirely of relations among observable quantities, such as frequencies and intensities of spectral lines. His theory, therefore, employs no atomic models but consists of mathematical relations between certain sets of numbers, or "matrices." Its mathematical processes are consequently those of the algebra of matrices.

Although appearing at first to have little in common, it turns out that the wave mechanics of de Broglie and Schroedinger and the quantum mechanics of Heisenberg lead to results which are the same and which are in better agreement with the results of experiment than have been those of any of our preceding theories.

It seems that we now have a way out of the difficulties presented by the dual nature of matter and energy, and are again proceeding along the alluring path of scientific adventure, with better equipment and better prospects of success than ever before in searching out the secret system in the processes of nature.

> The most outrageous ill fortune must in the end yield to the untiring courage of philosophy.
>
> —Edgar Allan Poe

APPENDIX

I. **Dimensional Reasoning.** Since only quantities of the same kind can be added, subtracted, and equated, it follows that the dimensional formulas of all terms of an equation must be the same. Consequently if one is uncertain of the exact form of an equation, it may be checked by dimensional reasoning.

Thus in Bernoulli's theorem (Eq. 130),

$$h + \frac{p^1}{D} + \frac{v^2}{2g} = H$$

the dimensional formulas for the individual terms are:

$$h = [L]$$

$$\frac{p^1}{D} = \frac{p}{gD} = \frac{[ML^{-1}T^{-2}]}{[LT^{-2}][ML^{-3}]} = \frac{[ML^{-1}T^{-2}]}{[ML^{-2}T^{-2}]} = [L]$$

$$\frac{v^2}{2g} = \frac{[LT^{-1}]^2}{[M^0L^0T^0][LT^{-2}]} = \frac{[L^2T^{-2}]}{[LT^{-2}]} = [L]$$

Hence each term of Eq. 130 has the dimensions of length, or height, and is called a "head"; and the equation is consistent.

The algebraic expression for a quantity may often be determined by dimensional reasoning. For example, suppose that we suspect that centrifugal force F depends upon the mass M of the body, its speed v, and the radius r of its path.

Then we may write $$F = M^x v^y r^z \tag{a}$$

where x, y, and z are to be determined.

Putting the dimensional formula for each quantity in Eq. (a),

$$[MLT^{-2}] = [M]^x[LT^{-1}]^y[L]^z$$
$$= [M^xL^{y+z}T^{-y}]$$

Since the two sides must be identical, we may equate the exponents of the M's, L's, and T's respectively.

Therefore, $$x = 1;\ y + z = 1;\ -y = -2$$

whence $$x = 1,\ y = 2,\ z = -1$$

Quantity
Length
Mass
Time
Force
Energy
Power
Resistance
Current
Potential
Charge
Capacitance
Inductance
Magnetic flux
Magnetic flux density
Magnetic field strength
Dielectric flux density
Dielectric field strength
Magnetic pole strength

* From Jauncey and Langsdorf,

and on putting these values back in Eq. (a),

$$F = M^1v^2r^{-1}$$

$$= \frac{Mv^2}{r} \quad \text{which is the expression sought.}$$

This method will not give the value of a numerical coefficient, however, because all pure numbers have the dimensional formula $[M^0L^0T^0]$ and drop out of the equations.

II. Conversion Table *

M.K.S.		C.G.S.
1 meter	equals	100 cm
1 kg		1000 gm
1 sec		1 sec
1 newton		10^5 dynes
1 joule		10^7 ergs
1 watt		10^7 ergs/sec

		Electrostatic Units		Electromagnetic Units
1 ohm	equals	$1/(9 \times 10^{11})$ statohm	equals	10^9 abohms
1 ampere		3×10^9 statamperes		10^{-1} abampere
1 volt		1/300 statvolt		10^8 abvolts
1 coulomb		3×10^9 statcoulombs		10^{-1} abcoulomb
1 farad		9×10^{11} statfarads		10^{-9} abfarad
1 henry		$1/(9 \times 10^{11})$ stathenry		10^9 abhenrys
1 weber		1/300 statweber		10^8 maxwells
1 weber/meter^2		$1/(3 \times 10^6)$ statweber/cm^2		10^4 gauss
1 M.K.S. oersted		3×10^7 statoersted		10^{-3} oersted
1 coulomb/4π meter^2		3×10^5 statcoulomb/4π. cm^2		10^{-5} abcoulomb/4π cm^2
1 volt/meter		$1/(3 \times 10^4)$ statvolt/cm		10^6 abvolts/cm
1 M.K.S. unit		1/300 statunit		10^8 abunits

M.K.S. Units and Dimensions, The Macmillan Company, 1940, by permission.

Natural Functions of Angles

Degrees	Radians	Sine	Tangent	Cosine
0	.0000	.0000	.0000	1.0000
1	.0175	.0175	.0175	.9998
2	.0349	.0349	.0349	.9994
3	.0524	.0523	.0524	.9986
4	.0698	.0698	.0699	.9976
5	.0873	.0872	.0875	.9962
6	.1047	.1045	.1051	.9945
7	.1222	.1219	.1228	.9925
8	.1396	.1392	.1405	.9903
9	.1571	.1564	.1584	.9877
10	.1745	.1736	.1763	.9848
11	.1920	.1908	.1944	.9816
12	.2094	.2079	.2126	.9781
13	.2269	.2250	.2309	.9744
14	.2443	.2419	.2493	.9703
15	.2618	.2588	.2679	.9659
16	.2793	.2756	.2867	.9613
17	.2967	.2924	.3057	.9563
18	.3142	.3090	.3249	.9511
19	.3316	.3256	.3443	.9455
20	.3491	.3420	.3640	.9397
21	.3665	.3584	.3839	.9336
22	.3840	.3746	.4040	.9272
23	.4014	.3907	.4245	.9205
24	.4189	.4067	.4452	.9135
25	.4363	.4226	.4663	.9063
26	.4538	.4384	.4877	.8988
27	.4712	.4540	.5095	.8910
28	.4887	.4695	.5317	.8829
29	.5061	.4848	.5543	.8746
30	.5236	.5000	.5774	.8660
31	.5411	.5150	.6009	.8572
32	.5585	.5299	.6249	.8480
33	.5760	.5446	.6494	.8387
34	.5934	.5592	.6745	.8290
35	.6109	.5736	.7002	.8192
36	.6283	.5878	.7265	.8090
37	.6458	.6018	.7536	.7986
38	.6632	.6157	.7813	.7880
39	.6807	.6293	.8098	.7771
40	.6981	.6428	.8391	.7660
41	.7156	.6561	.8693	.7547
42	.7330	.6691	.9004	.7431
43	.7505	.6820	.9325	.7314
44	.7679	.6947	.9657	.7193
45	.7854	.7071	1.0000	.7071

SPECIFIC GRAVITY OF WATER AND MERCURY

Temperature 0°C	Water	Mercury
− 10	0.99815	13.6202
− 5	0.99930	13.6078
0	0.99987	13.5955
+ 1	0.99993	13.5930
2	0.99997	13.5906
3	0.99999	13.5881
4	1.00000	13.5856
5	0.99999	13.5832
6	0.99997	13.5807
7	0.99993	13.5782
8	0.99988	13.5758
9	0.99981	13.5733
10	0.99973	13.5708
11	0.99963	13.5684
12	0.99952	13.5659
13	0.99940	13.5634
14	0.99927	13.5610
15	0.99913	13.5585
16	0.99897	13.5561
17	0.99880	13.5536
18	0.99862	13.5512
19	0.99843	13.5487
20	0.99823	13.5462
21	0.99802	13.5438
22	0.99780	13.5413
23	0.99756	13.5389
24	0.99732	13.5364
25	0.99707	13.5340
26	0.99681	13.5315
27	0.99654	13.5291
28	0.99626	13.5266
29	0.99597	13.5242
30	0.99567	13.5217

III. Law of the Magnetic Circuit. This law plays the same important role in magnetic circuit problems that Ohm's law plays in problems of the electric circuit. It may be derived as follows for the simple case of a toroidal core, but its validity is general.

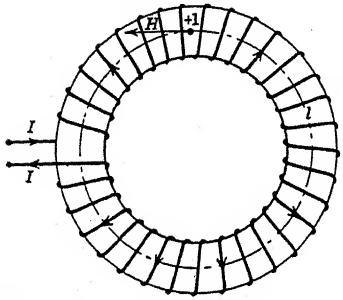


FIG. 639.

Consider a uniform coil wound on a toroidal core whose permeability is μ, and let the current in the coil be I abamperes (Fig. 639).

If H is the average magnetic field intensity for any radial cross-section A of the core, then the average flux density will be

$$B = \mu H \text{ lines/cm}^2\text{, or gausses,}$$

and the total flux Φ in the core will be

$$\Phi = BA = \mu HA \text{ lines, or maxwells}$$

whence $$H = \frac{\Phi}{\mu A} \text{ oersteds.}$$

Because of the complete symmetry of the core and coil, this H is the same at all points along the center line. Calling the length of the center line of the torus l, the magnetomotive force, or work to move a $(+ 1)$ magnetic pole around the magnetic circuit, is

$$\mathfrak{M} = Hl = \frac{\Phi}{\mu A} l$$

from which $$\Phi = \frac{\mathfrak{M}}{\frac{l}{\mu A}}. \tag{474}$$

This quantity $\frac{l}{\mu A}$ is called the reluctance ($\equiv \mathfrak{R}$) of the magnetic circuit, and its electromagnetic unit is here called the **rowland.** Equation 474, which may then be written,

$$\Phi_{\text{maxwells}} = \frac{\mathfrak{M} \text{ gilberts}}{\mathfrak{R} \text{ rowlands}} \tag{475}$$

is **Rowland's law.***

It will be seen that Rowland's law is similar to Ohm's law: magnetomotive force corresponds to electromotive force; reluctance, to resistance; and magnetic flux to electric current. But

* The law is attributed also to Bosanquet.

it should be borne in mind that magnetism does not flow whereas electricity does.

As one would expect, where a magnetic circuit is divided the sum of the fluxes in the branches equals the flux in the undivided part. This relation is mathematically analogous to the law of continuity of current for electricity (Sec. 406).

Since Kirchhoff's laws (Sec. 442) were derived from Ohm's law and the law of continuity of current, it follows that by substituting magnetic terms for the corresponding electric terms, Kirchhoff's laws so modified may be applied to the general solution of magnetic circuits. Expressed in symbols, the laws would be

1. At any junction, $\Sigma\Phi = 0.$
2. Around any closed circuit, $\Sigma\mathcal{M} = \Sigma\mathcal{R}\Phi.$

IV. **Reluctance.** Reluctance is the opposition that a magnetic path offers to the establishment of a magnetic flux in it. From III above it is measured by the relation,

$$\mathcal{R} \equiv \frac{l}{\mu A}$$

where $\mathcal{R}$ is unity when l is 1 cm, A is 1 cm², and μ is 1.

Hence this electromagnetic unit of reluctance, **the rowland,* is defined as the reluctance of a magnetic path 1 cm long, having a cross section of 1 cm² and the permeability of a vacuum.**

Since Rowland's law corresponds to Ohm's law which was the basis for the derivation of Equations 322 and 323, precisely similar relations may be deduced for reluctances:

1. Reluctances in series:

$$\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3 + \cdots \tag{476}$$

2. Reluctances in parallel:

$$\frac{1}{\mathcal{R}} = \frac{1}{\mathcal{R}_1} + \frac{1}{\mathcal{R}_2} + \frac{1}{\mathcal{R}_3} + \cdots \tag{477}$$

where $\mathcal{R}$ is the combined reluctance and $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$, $\cdots$ are the individual reluctances.

* This unit has not been adopted officially.

INDEX